

ABSTRACT

Title:

A HYBRID METHODOLOGY FOR
MODELING RISK OF ADVERSE EVENTS
IN COMPLEX HEALTHCARE SETTINGS

Reza Kazemi- Tabriz, Ph.D., 2011

Directed by:

Professor Ali Mosleh, Reliability Engineering
Program

Assistant Professor Meghan Dierks, Harvard
Medical School

Despite efforts to provide safe, effective medical care, adverse events still occur with some regularity. While risk cannot be entirely eliminated from healthcare activities, an important goal is to develop effective and durable mitigation strategies to render the system 'safer'. In order to do this, though, we must develop models that comprehensively and realistically characterize the risk. In the healthcare domain, this can be extremely challenging due to the wide variability in the way that healthcare processes and interventions are executed and also due to the dynamic nature of risk in this particular domain. In this study we have developed a generic methodology for evaluating dynamic changes in adverse event risk in acute care hospitals as a function of organizational and non-organizational factors, using a combination of modeling formalisms. First, a system dynamics (SD) framework is used to demonstrate how organizational level and policy level contributions to risk evolve over time, and how policies and decisions may affect the general system-level contribution to adverse event risk. It also captures the feedback of organizational factors and decisions over time and the non-linearities in these feedback effects. Second, Bayesian Belief Network (BBN) framework is used to represent patient-level factors and also physician level decisions and factors in the management of an individual patient, which contribute to the risk of hospital-acquired adverse event. The model is intended to support hospital decisions with regards to staffing, length of stay, and investment in safeties, which evolve dynamically over time. The methodology has been applied in modeling the two types of common adverse events; *pressure ulcers* and *vascular catheter-associated infection*, and has been validated with eight years of clinical data.

A HYBRID METHODOLOGY FOR MODELING RISK OF ADVERSE EVENTS
IN COMPLEX HEALTHCARE SETTINGS

By

Reza Kazemi –Tabriz

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
Of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Professor Ali Mosleh, Chair

Assistant Professor (Harvard Medical School) Meghan Dierks, Co-Chair

Professor Mohammad Modarres

Professor Marvin Roush

Professor Gregory Baecher

Professor Robert Sprinkle (Dean's Representative)

©Copyright by

Reza Kazemi-Tabriz

2011

to my parents

Acknowledgements

I would like to express my deepest gratitude to my advisor and mentor, Dr. Ali Mosleh for his guidance, support, patience and encouragement during my graduate studies and this research. I am forever grateful to him and all he has taught me, and can never thank him enough for his kindness.

My sincere thanks are due to Dr. Meghan Dierks who took so much time out of her busy schedule to help and guide me throughout this research. Without her, this thesis would have not been realized.

Thanks go to the panel of experts for their invaluable input and insightful remarks in the course of this research.

I would also like to thank Professors Modarres, Roush, Baecher and Sprinkle who served on my doctoral committee.

My gratitude is extended to my parents, my sisters and Ladan, for their constant encouragement, love and support.

Table of Contents

ABSTRACT	1
Table of Contents.....	iv
List of Tables.....	vii
List of Figures.....	viii
Abbreviations.....	xiii
1 Introduction	1
1.1 Motivation	1
1.2 Significance of the Problem	2
1.3 Research Objective.....	4
1.4 Summary of Approach	4
1.5 Dissertation Outline.....	6
2 Related Work.....	7
2.1 Informal Risk Analysis Methods.....	8
2.1.1 Failure Mode and Effect Analysis (FMEA) in Engineering	8
2.1.2 FMEA in Healthcare	8
2.1.3 Shortcomings of FMEA in Medical Applications	9
2.1.4 Miscellaneous Approaches to Risk Assessment.....	12
2.2 Formal Risk Analysis Methods	13
2.2.1 PRA in Engineering.....	14
2.2.2 PRA in Healthcare.....	16
2.2.3 Shortcomings of PRA in Medical Applications	16
2.2.4 HRA in Engineering	17
2.2.5 HRA in Healthcare	17
2.2.6 Shortcomings of HRA in Healthcare	21
3 Methodology.....	23
3.1 Overview	23
3.2 Bayesian Belief Networks (BBN)	24
3.2.1 Introduction.....	24
3.2.2 BBN Elements.....	26
3.2.3 Construction and Quantification of BBNs.....	31
3.3 System Dynamics (SD)	40
3.3.1 Introduction	40
3.3.2 Building Blocks.....	41

3.4	SD/BBN Combination.....	45
3.5	Information Sources for Quantification of Models	47
3.5.1	<i>Data form Actual Operating Experience.....</i>	47
3.5.2	<i>Adverse Event Data.....</i>	48
3.5.3	<i>Expert Elicitation.....</i>	50
3.5.4	<i>Who is an Expert?.....</i>	51
3.5.5	<i>Elicitation Methods.....</i>	53
3.5.6	<i>The Panel of the Experts.....</i>	56
3.5.7	<i>Eliciting the Structure of the Models; Qualitative Part.....</i>	57
3.5.8	<i>Eliciting the Parameters of the Model (Model Quantification).....</i>	62
3.5.9	<i>Formats of Elicitation.....</i>	62
3.5.10	<i>Challenges and Generic Issues in Eliciting Expert Opinion.....</i>	63
3.6	Validation and Verification.....	65
4	Model Development for Adverse Events; Pressure Ulcer	68
4.1	Introduction.....	68
4.1.1	<i>Risk Assessment Tools</i>	70
4.2	Pressure Ulcer BBN Development.....	77
4.3	Pressure Ulcer BBN Quantification.....	83
4.3.1	<i>Modeling Techniques.....</i>	84
4.3.2	<i>The Quantification Process.....</i>	88
4.3.3	<i>The Conditional Probability Table (CPT).....</i>	93
4.4	Pressure Ulcer BBN Validation	114
4.4.1	<i>Qualitative Validation of Pressure Ulcer BBN.....</i>	116
4.4.2	<i>Quantitative Validation of Pressure Ulcer BBN</i>	118
5	Model Development for Adverse Events; Vascular Catheter-Associated Infection	125
5.1	Introduction.....	125
5.2	Vascular Catheter-Associated Infection BBN Development.....	126
5.3	Vascular Catheter-Associated Infection BBN Quantification	130
5.4	Vascular Catheter-Associated Infection BBN Validation	138
5.4.1	<i>Vascular Catheter-Associated Infection BBN Qualitative Validation</i>	138
5.4.2	<i>Vascular Catheter-Associated Infection BBN Quantitative Validation.....</i>	138
6	Dynamic Model of Hospital-level Factors Affecting the Risk of Adverse Events.....	140
6.1	Dynamic Model Development.....	141
6.1.1	<i>Introduction</i>	141
6.1.2	<i>Model Developing Process.....</i>	141
6.1.3	<i>The System Dynamics Model.....</i>	143
6.2	Dynamic Model's Quantification.....	152
6.2.1	<i>Soft Factors in the Model: "Pressures"</i>	153

6.2.2	<i>Data Driven Factors</i>	172
6.2.3	<i>Magnitude of the Effects of Cost Cutting Strategies on the Risk of Adverse Events</i>	179
6.2.4	<i>Other Constants of the Model</i>	198
6.3	<i>Dynamic Model’s Validation</i>	202
6.3.1	<i>Informal / Qualitative Model Validation</i>	203
6.3.2	<i>Formal/ Quantitative Model Validation</i>	206
7	<i>The Hybrid Model; Analysis and Results</i>	212
7.1	<i>Evaluation and Validation of the Hybrid Model</i>	212
7.1.1	<i>Hybrid Model Performance; Risk of Pressure Ulcer</i>	213
7.1.2	<i>Hybrid Model Performance; Risk of Line Infection</i>	216
7.1.3	<i>Hybrid Model Performance; Risk of Line Infection and Pressure Ulcer</i>	219
7.2	<i>Use of the Risk Model; Hypothetical Examples</i>	223
7.3	<i>Uncertainty Analysis</i>	228
7.3.1	<i>Uncertainty Analysis; Hybrid Model for Risk of Pressure Ulcer</i>	228
7.3.2	<i>Uncertainty Analysis; hybrid Model for Risk of Line Infection</i>	230
7.3.3	<i>Uncertainty Analysis; hybrid Model for Risk of Line Infection and Pressure Ulcer Combined</i>	233
7.4	<i>Importance Measures</i>	234
7.4.1	<i>Risk Reduction Worth</i>	234
7.4.2	<i>Dependencies in Risk Importance Measure’s Quantification</i>	235
7.4.3	<i>Example; Importance Measure for Pressure to Optimize LOS in the Risk of Line Infection</i>	236
7.5	<i>Model Requirements for Application</i>	239
8	<i>Summary and Conclusion</i>	241
8.1	<i>Challenges</i>	242
8.2	<i>Contributions</i>	243
8.3	<i>Future Work</i>	244
	<i>Appendix A. Model Validation Interview Guide, Pressure Ulcer BBN</i>	245
	<i>Appendix B. Model Validation Interview Guide, line Infection BBN</i>	250
	<i>Appendix C. Model Validation Interview Guide, System Dynamics Model; Round 1&2</i>	254
	<i>Appendix D. Model Validation Interview Guide, System Dynamics Model; Round 3</i>	264
	<i>References</i>	270

List of Tables

Table 2-1. Categories of HRA techniques (Lyons, 2004)	19
Table 3-1. List of variables and their states	27
Table 3-2. Hospital acquired adverse events	50
Table 3-3. Panel of experts	57
Table 4-1. Different stages of pressure ulcer	69
Table 4-2. Expert opinion; staff adequacy	98
Table 4-3. Expert opinion; probability of using assistive devices, given staff adequacy	103
Table 4-4. Expert opinion; conditional probability of adequacy of frequency of move	107
Table 4-5. Expert opinion; probability of pressure ulcer due to frequency of move (left)	111
Table 4-6. Expert opinion; probability of pressure ulcer due to frequency of move (right)	111
Table 4-7. Probabilities of the effect of risk factors, independently, on pressure ulcer	113
Table 4-8. Expert opinion; qualitative evaluation of pressure ulcer BBN	117
Table 4-9. Relative frequency of hospital acquired pressure ulcer; 2003 -2011	118
Table 5-1. Relative frequency of hospital acquired bloodstream infection; 2002-2009	137
Table 5-2. Expert Opinion; qualitative evaluation of line infection BBN	138
Table 6-1. Fiscal data	152
Table 6-2. Operating margin and safety investment data, 2003-2011	169
Table 6-3. Probability of LOS too short to meet all patient's needs	173
Table 6-4. Sample data records extracted to indicate whether the patient had prolonged LOS ...	178
Table 6-5. Expert opinion; level of pressure to optimize LOS triggering early discharge	180
Table 6-6. Magnitude of change in the risk of hospital acquired adverse events due to staffing, prolonged and shortened LOS	182
Table 6-7. Expert opinion; level of pressure to optimize LOS triggering prolonged LOS	183
Table 6-8. Case-mix complexity distribution across hospital; descriptive statistics	187
Table 6-9. Experts' opinion; combination of level of pressure to reduce operational cost, and complexity score effecting the probability of adverse events due to understaffing	190
Table 6-10. Safety investment data and prevalence of pressure ulcer	193
Table 6-11. Experts' opinion; level of ability/willingness to invest in safety and the effect of risk of adverse events	195
Table 6-12. Experts' opinion on magnitude of change in the risk of adverse events due to lack of investment in safety; aggregated	196
Table 6-13. Experts' opinion; magnitude of change in the risk of adverse events due to lack of investment in safety	197
Table 6-14. Average LOS for pressure ulcer and non-pressure ulcer cohort	200
Table 6-15. Expert opinion; qualitative evaluation of system dynamics module (organizational level contributions to risk)	205
Table 7-1. Hybrid models' error in predicting the probability of pressure ulcer for each year	215
Table 7-2. Hybrid models' error in predicting the probability of line infection for each year	218

Table 7-3..Hybrid models’ error in predicting the probability of line infection for each year, calibrating with both BBNs	222
Table 7-4.Importance measures of pressure functions in the model for years 2002-2009.....	237

List of Figures

Figure 2-1.Sample form of the completed FMEA analysis; blood transfusion process (Burgmeier, 2002).....	11
Figure 2-2.Classical PRA methodology (Figure originally composed by Futron corporation, NASA contractor for ISS PRA)	15
Figure 3-1.A simple belief network; event X (parent node, cause node), causes/influences event Y (child node, effect node).....	24
Figure 3-3.Lung cancer BBN	28
Figure 3-4.Overall causal structure of BBN (Kjaerulff and Madsen, 2008)	33
Figure 3-5.Balancing and reinforcing loops	43
Figure 3-6: Stock and flow diagram.....	44
Figure 3-7.SD and BBN hybrid modeling environment.....	46
Figure 3-8.Model construction and qualitative validation process.....	61
Figure 3-9.Simulation Model Verification and Validation in the Modeling Process (Sargent, 2001).....	66
Figure 4-1.Pressure ulcer risk factors (Keller, 2002)	72
Figure 4-2.Barden scale risk assessment worksheet.....	74
Figure 4-3.Pressure ulcer BBN.....	79
Figure 4-4.Parent divorcing.....	86
Figure 4-5.Pressure ulcer BBN, transformation 1	91
Figure 4-6.Pressure ulcer BBN, transformation 2	92
Figure 4-7.Pressure ulcer BBN, final transformation.....	93
Figure 4-8.Joint distribution of the parameters, of the distribution of “Probability of In-Adequate Staffing”	101
Figure 4-9.Average distribution of the population variability distribution set.....	101
Figure 4-10.Average distribution of the population variability distribution set for “probability of assistive devices NOT used, given adequate staffing”	104
Figure 4-11.Average distribution of the population variability distribution set for “probability of assistive devices NOT used, given in-adequate staffing”	105
Figure 4-12.Average distribution of the population variability distribution set for probability of in-adequate frequency of move (a).....	107
Figure 4-13.Average distribution of the population variability distribution set for probability of in-adequate frequency of move (b).....	108
Figure 4-14.Average distribution of the population variability distribution set for probability of in-adequate frequency of move (c).....	109

Figure 4-15.Average distribution of the population variability distribution set for probability of in-adequate frequency of move (d).....	110
Figure 4-16.Average distribution of the population variability distribution for effect of frequency of movement on pressure ulcer.....	112
Figure 4-17.Average distribution of the population variability distribution for effect of using assistive devices on pressure ulcer	113
Figure 4-18.Probability of hospital acquired pressure ulcer.....	114
Figure 4-19.Distribution of the probability of hospital acquired pressure ulcer	118
Figure 4-20.Posterior Distribution of Probability of Pressure Ulcer; BBN Predication Adjusted Using Bayesian Model Uncertainty Method	124
Figure 5-1.Vascular catheter related infection BBN	126
Figure 5-2.Line infection BBN; transformed	132
Figure 5-3.Sample data records used for line infection BBN quantification	132
Figure 5-4.Distribution of line infection probability; 2002-2009.....	139
Figure 6-1.Revenue deficit-loss of service downward spiral	146
Figure 6-2.System dynamics module; highest level of abstraction	147
Figure 6-3.System dynamics module; detailed version.....	150
Figure 6-5.Operating margin-pressure to close revenue gap relationship.....	156
Figure 6-6.Operating margin-pressure to close revenue gap relationship; format options	157
Figure 6-7.Expert opinion; operating margin/pressure to reduce revenue gap relationship.....	157
Figure 6-8.Operating margin/pressure to reduce revenue gap relationship; functional form.....	159
Figure 6-9.Pressure to close revenue gap-pressure to optimize LOS relationship; format options	161
Figure 6-10.Expert opinion; pressure to reduce revenue gap relationship/pressure to optimize LOS relationship.....	161
Figure 6-11.Pressure to reduce revenue gap relationship/pressure to optimize LOS relationship; functional form.....	163
Figure 6-12.Pressure to close revenue gap/pressure to reduce operational costs relationship; format options.....	166
Figure 6-13.Expert opinion; Pressure to close revenue gap/pressure to reduce operational costs relationship	166
Figure 6-14.Pressure to close revenue gap/pressure to reduce operational costs relationship; functional form	167
Figure 6-15.Operating margin and safety investment; regression analysis.....	169
Figure 6-16.Pressure to close revenue gap/willingness ability to invest in safety; format options	170
Figure 6-17.Expert opinion; pressure to close revenue gap/willingness ability to invest in safety relationship	171
Figure 6-18.Number of days in inappropriate early discharge; pdf.....	174
Figure 6-19.Number of days in inappropriate early discharge; cdf.....	175
Figure 6-20.Experts' opinion on level of pressure to optimize LOS triggering early discharge; aggregated	181
Figure 6-21.Experts' opinion on level of pressure to optimize LOS triggering prolonged LOS; aggregated	184

Figure 6-22. Case-mix complexity distribution across hospital	187
Figure 6-23. Calculating the probability of understaffing	188
Figure 6-24. Case-mix complexity distribution across hospital; CDF	192
Figure 6-25. Experts' opinion on level of ability/willingness to invest in safety and the effect of risk of adverse events; aggregated.....	195
Figure 6-26. Distribution of increased LOS due to pressure ulcer	200
Figure 6-27. System dynamics model validation framework.....	206
Figure 6-28. Extreme case testing; pressure variables in the model when pressure to close revenue gap is at maximum	209
Figure 6-29. Extreme case testing; risk of pressure ulcer when pressure to close revenue gap is at maximum increases 80%	209
Figure 7-1. Risk of hospital acquired pressure ulcer; hybrid model predictions for each year versus actual data.....	215
Figure 7-2. Risk of line infection; hybrid model predictions for each year versus actual data	218
Figure 7-3. Risk of PU; hybrid model predictions for each year versus actual data calibrating the model using both BBNs	221
Figure 7-4. Risk of line infection; hybrid model predictions for each year versus actual data calibrating the model using both BBNs.....	222
Figure 7-5. Setting evidence in line infection BBN	224
Figure 7-6. Compiling BBN to update the probability of line infection	225
Figure 7-7. Predicting patient a's probability of line infection, influenced by organizational factors	225
Figure 7-8. Decline in the risk of line infection over the next 10 years, due to increase in operating margin by 0.2% per year	226
Figure 7-9. Increase in the risk of line infection over the next 10 years, due to decrease in operating margin by 0.2% per year	227
Figure 7-10. Confidence bounds for model predictions on risk of pressure ulcer as a result of uncertainty over baseline risk of pressure ulcer	229
Figure 7-11. Risk of pressure ulcer from clinical data	229
Figure 7-12. Confidence bounds for model predictions on risk of pressure ulcer as a result of uncertainty over baseline risk of pressure ulcer and uncertainty over the probability of prolonged LOS	230
Figure 7-13. Confidence bounds for model predictions on risk of line infection as a result of uncertainty over baseline risk of line infection	231
Figure 7-14. Risk of line infection from clinical data	231
Figure 7-15. Confidence bounds for model predictions on risk of line infection as a result of uncertainty over baseline risk of pressure ulcer and uncertainty over the probability of prolonged LOS	232
Figure 7-16. Confidence bounds for model predictions on risk of line infection as a result of uncertainty over baseline risk of line infection; the case of the hybrid model with both line infection and pressure ulcer BBNs	233
Figure 7-17. Importance measures of model's pressure functions over time.....	238

Abbreviations

BBN	Bayesian Belief Networks
CNS	Central Nervous System
ESD	Event Sequence Diagram
FMEA	Failure Mode and Effect Analysis
FT	Fault Tree
FTA	Fault Tree Analysis
HRA	Human Reliability Analysis
PRA	Probabilistic Risk Analysis
PU	Pressure Ulcer
PVD	Peripheral vascular Disease
RAS	Risk Assessment Scale
RIM	Risk Importance Measure
RRW	Risk Achievement Worth
SD	System Dynamics

1 Introduction

1.1 Motivation

Despite efforts to provide safe, effective medical care, adverse events still occur with some regularity. While risk cannot be entirely eliminated from healthcare activities, our goal is to develop effective and durable mitigation strategies to render the system ‘safer’. In order to do this, though, we must develop models that comprehensively and realistically characterize the risk. In the healthcare domain, this can be challenging for a number of reasons. In contrast to traditional engineering domains, there can be wide variability in the way that healthcare processes and interventions are executed. This variability is due not only to organizational and human performance issues, but also to the high degrees of uncertainty associated with management of most clinical conditions, and variability in the quality/reliability of the information used to make decisions. Another modeling challenge is the dynamic nature of risk in this particular domain. Characteristics or conditions of the clinical care environment that might pose a hazard can change as a function of time, and/or the changing state of internal and external factors. Also, for individual patients, exposure to these internal and external factors varies as a function of time and underlying medical condition.

The goal of this research is to develop a generic methodology for evaluating dynamic changes in adverse event risk in acute care hospitals as a function of organizational factors and non-organizational factors, using a combination of modeling formalisms. First, a system dynamics (SD) framework will be used to capture changes in the level of

risk as a function of: duration of hospital stay, complexity of the patient's condition, Financial wellbeing of the organization, policies and decisions taken to respond to the financial standing of the organization and constraints imposed by external agencies (e.g., insurers and regulatory/certification authorities) on operational decisions.

The SD framework enables us to capture feedback reinforcement of specific factors over time, and non-linearities in these effects. This would not be possible using conventional risk analytic techniques. Second, Bayesian methods will be applied to provide input to some of the variable nodes. The Bayesian Belief network is used to capture patient level factors and conditions and patient-provider level factors which correspond to provider's decisions in treating a patient and patient's responses to such interventions. Using nine years' of clinical data and domain expertise from one of Harvard Medical Schools major teaching hospitals, we will also validate the performance of this methodology in studying this problem.

1.2 Significance of the Problem

In spite of increased attention to quality, and efforts to provide safe medical care, adverse outcomes are still frequent in clinical practice (Leape, 94). Reports from various sources indicate that a substantial number of hospitalized patients suffer treatment caused injuries, while in the hospital (Leape, 97). Harvard Medical Practice Study in 1991 (Brennan, et.al., 1991), based on an study of injuries in patients in the state of New York in 1984, reported that nearly 4% of all hospitalized patients suffered injuries that prolonged their length of stay in hospital, or resulted in some level of disability, and 14% of these injuries were fatal. Assuming a homogeneous population, with extrapolating, 1.3 million people are harmed and 180,000 people die in United States only at least in part

because of an injury during their hospitalization. Moreover, it was found that 69% of those injuries were due to errors and therefore preventable (Leape, et.al., 1991). Other studies have reported different statistics. The Institute of Medicine (IOM) reported in 1999 that 44,000 to 98,000 people die in hospitals each year as the result of medical errors (Kohn et al, 2000). This exceeds combined toll from motorcycle crashes, suicides, falls, poisoning and drowning. The report also indicates that medical error costs the nation \$37.6 billion each year where about \$17 billion of those costs are associated with preventable errors.

Before the IOM's report in 1991, formal approaches to the analysis of adverse events were relatively uncommon. Attention attracted to a number of highly publicized medical error cases that resulted in death or injury of patients (Bogner, 2001) combined with the realization of the fact that more could be done in hospitals to prevent injuries due to errors, led to a significant increase both in investigating the causes of error and finding effective error prevention methods.

Despite the magnitude of the problem, current analysis of adverse events in healthcare settings continues emphasis on individual case studies. Efforts to understand the nature of aggregate risk through formal methods have been limited. The lack of a comprehensive modeling formalism that is able to demonstrate the causal relationship between the factors effecting risk of adverse event and how this risk might evolve in time under the influence of organizational, individual and policy level factors, has been the major rationale for the currently proposed research.

1.3 Research Objective

The broad goal of this research is to develop and apply a methodology to evaluate dynamic changes in adverse event risk in acute care hospitals. Although some of the risk is related to the underlying complexity of care and severity of illness in the patient population, a significant portion may be related to the structure of the system – most notably, the operational policies, incentive structures, and constraints imposed by third-parties who finance care. Any efforts to redesign the system, however, must be preceded by careful modeling and analysis to demonstrate exactly how the policies and features of the system influence risk. In this research, we attempt to build models that demonstrate these system-level influences and how they dynamically shape risk in the healthcare domain.

1.4 Summary of Approach

In order to make these models both realistic and useful (i.e., capable of providing new insight), we have adopted a hybrid modeling strategy that incorporates both system dynamics (SD) principles and Bayesian belief networks (BBN). The SD-BBN combination enables us to capture some of the more important features of the healthcare environment.

Input to the quantitative component of the model will be derived from clinical data and information from domain experts. The model will use Bayesian data analysis techniques and SD-BBN framework to integrate different types of data.

The major phases in this research are:

- 1- Qualitative modeling
- 2- Data collection
- 3- Quantification and calibration
- 4- Validation

The model building process started with developing a qualitative understanding of some of the major risk scenarios, with an adverse event as the end state and some organizational level decision or policy as the initiating event.

In data collection phase, the goal is to identify the type and quality of data available and to select data relevant to the factors in the model. This mainly involves studying the cases of adverse events and relating them to nodes in the model. Eight years years' of clinical data from one of Harvard Medical School's major teaching hospitals is available at this phase.

Having built the qualitative model in a System Dynamic framework, the model is tested calibrated and fine-tuned with data obtained in phase 2. For this stage of the process, six year worth of data is used (from the available 8 years). At the last phase, the results of the model are validated against data from the remaining years.

1.5 Dissertation Outline

Chapter 2 of this dissertation, reviews current risk assessment and risk analysis literature in healthcare domain, and highlights the fact that there is need for more comprehensive and realistic modeling and representation of risk in this domain, in order to provide insight to the decisions to be made and policies to be set. The hybrid methodology and its components, Bayesian Belief Networks (BBN) and System Dynamics (SD) modeling formalisms, have been described in chapter 3. Chapter 3 also discussed the sources of information (i.e. clinical data and expert opinion) used in this research.

In this research, we have developed two BBN models for the risk of two specific adverse events; pressure Ulcer and Vascular catheter-Associated Infection (i.e. line infection). We have also developed a system dynamics model to represent the organizational-level contributions to risk of adverse events. Chapter 4 through 6 describes these models, their development process, their quantification and their validation. Chapter 7 discusses the hybrid model that consists of the combination of the system dynamic module (for organizational level factors) and the BBNs for the 2 adverse events (pressure ulcer and line infection). In chapter 7 we also present a set of uncertainty analysis performed on the hybrid model. This chapter also contains a discussion on risk importance measures we have developed for the hybrid, dynamic model.

Finally chapter 8, addresses the contributions of this research and also the potential future work needed to further improve the application of this approach as an decision making tool.

2 Related Work

As noted above, prior to the publication of the IOM report in 1991, formal approaches to the analysis of adverse events were relatively uncommon. In response to this, some quasi-regulatory authorities (notably the Joint Commission, an independent body that accredits hospital and other healthcare facilities), and some Federally-sponsored research organizations (e.g., the Department of Health and Human Services' Agency for Healthcare Research and Quality) have encouraged use of some modeling formalisms originally developed for the engineering discipline and non-healthcare disciplines. In most cases, their application to healthcare either has been experimental in nature, or informal. We have categorized these approaches, into two categories; Formal and Informal risk analysis methods. Generally informal risk analysis methods, such as failure mode and effect analysis; a) lack a systemic view and perspective compared to the formal methods, such as probabilistic risk assessment methods, b) are mostly, essentially qualitative, and c) lack an explicit causal perspective.

This section reviews these approaches and their application in healthcare and discusses the shortcomings of each of these methods in addressing risk of adverse event in medical domain.

2.1 Informal Risk Analysis Methods

2.1.1 Failure Mode and Effect Analysis (FMEA) in Engineering

Failure mode and effect analysis (FMEA) examines high risk processes to identify required improvements to reduce the probability of adverse events. It has been used in industry (e.g. manufacturing, aviation) for over 30 years to assess system safety.

The FMEA procedure is well documented in the military handbook (MIL-HDBK-338 and MIL-HDBK-338B) as a military standard. FMEA is done in two phases. The first phase is the identification of the potential failure modes and their effects. The second phase is performing criticality analyses to determine the severity of failure modes identified in phase one.

2.1.2 FMEA in Healthcare

FMEA is perhaps the most popular engineering risk analysis methods used by healthcare organizations. Its use was promoted by the Joint Commission, an independent organization that accredits hospitals and healthcare delivery organizations. As a condition of accreditation, healthcare organizations are required by the Joint Commission to conduct at least one FMEA annually on a healthcare process. The Veterans Health Administration has promoted the use of this technique, and has developed a modified version of the traditional industrial/military FMEA that emphasizes qualitative analysis of healthcare processes (DeRosier et.al. 2002). Published studies demonstrating the use of FMEA in healthcare include, an FMEA for reducing risk in blood transfusion (Burgmeier, 2002), FMEA in improving a drug distribution system (McNally, et.al.

1997), drug prescribing process (Saizy-Callaert, et.al., 2001) and intravenous drug infusion (Apkon et.al. 2004) and application in safety improvement of the production of pediatric parenteral nutrition solutions (Bonnabry, et.al., 2005). Figure 2-1 is a sample form of the completed FMEA analysis for blood transfusion process (Burgmeier, 2002).

2.1.3 Shortcomings of FMEA in Medical Applications

To date, most healthcare organizations have found that much of the utility of an FMEA lies in having healthcare professionals gather and map out the medical processes and procedures. The FMEA process brings together a multidisciplinary team, creating an opportunity for different types of providers to understand parts of a clinical process that without which they may not have been aware. This qualitative process modeling often results in a broader understanding (at the organizational level) of the dependencies and vulnerabilities of the healthcare process being modeled, and through this, probably contributes to some degree of risk management. FMEA activities in healthcare are rarely quantitative. Also, the FMEA process is highly subjective and dependent on the experience level of the analyst, and may not capture many potential failures. Shebl, Franklin and Barber (2009), conduct a study to test the reliability of FMEA analysis within a hospital setting, by recruiting two teams to conduct separate FMEAs in parallel on the same topic by following the basic FMEA steps including mapping the process of care, identifying potential failures of the process, determining the severity, probability and detectability scores for the failures and making recommendations to decrease the detected failures. The results indicated that even though each group identified 50 failures only 17% of them were common to both teams, and due to different severity, detectability and risk scores, the prioritization of failures were different. They conclude that these

discrepancies make it impossible to reliably identify failures that are to be prioritized, and optimally allocate resources, time, effort and money to improve patient safety.

FMEA might be a useful tool to investigate a particular risk, but is completely ineffective in identifying and describing how policies and decisions which are dominant contributors in healthcare, influence risk.

PFMEA: Blood Transfusion

(1) Process Step #1: Blood order entered

(2) Date: 2/27/01

(3) Description: Order for T & Xmatch/blood/blood product. Two people visualize order, screen, and print out right patient/right blood.

Failure Mode	Cause of Failure	Effects of Failure	Ooc	SV	DT	RPN	Design Action (Solutions)	Design Validation
Two people do not always visualize "order entry" for blood products (often not available). 1.1	1.11—Immediate patient care elsewhere is more important than entering order (when given a choice). 1.12—Nurses do not fully understand the consequences of a decision <u>not</u> to enter an order & instead address an immediate patient need elsewhere. 1.13—Nurses do not fully utilize one another as resources (i.e. make a decision for someone else that they are too busy to help). 1.14—Order entrant (eg, clerk, nurse) finds it more rewarding to expediently "get things done" without entering order correctly because it is reinforcing to get things done vs correctly and there are no manager consequences because of unawareness or conflict avoidance. 1.15—Current policy is not explicit that 2 people must visualize order is correct before sent.	<ul style="list-style-type: none"> Wrong patient gets "suck" Waste personnel hours & resources (eg, supplies) Delay in treatment to appropriate patient The up scarce blood resources Increase patient's level of risk Increase patient's length of stay Demoralize person not involved (eg, "I could have harmed/killed someone") 	10	3	9	270	<p>1.111</p> <ul style="list-style-type: none"> Blood-specific order form used by all departments and on computer screen that is completed by physician (RN completes telephone orders). Training on use of form will occur for everyone participating in blood transfusion process. Order form faxed to BTS and double-checked against computer entry. Variance plan needs to be developed for OR. Policy revision reflects this process. 1.112—Physician will enter blood order in computer (long-term). 	<p>1.1111—Data are collected on variances of paper and computer entries. Incidents of documented variance are reported to manager and quarterly to QIC/NAC. 1.1112—Manager will confront employees who are associated with variance to correct performance.</p>
Blood order is interpreted differently from the physician's intent. 1.2	1.21—Inconsistent physician language used when ordering blood. 1.22—RN/unit secretary reluctant to verify order with physician because they'll be yelled at. 1.23—Physicians are not themselves directly entering order into computer.	<ul style="list-style-type: none"> Same as bullet points 2-6 above Reinforces codependent behavior of RN/secretary and inappropriate behavior of physicians who do yell 	5	8	6	240	<p>1.21—Above solution addresses this problem, too.</p>	

* T, type; X, cross; Ooc, occurrence; SV, severity; DT, detectability; RPN, risk priority number; RN, registered nurse; BTS, blood transfusion service; OR, operating room; QIC, Quality Improvement Committee; NAC, Nursing Administrative Committee.
Format source: Arnan JM, Gyna FM: Juran's Quality Research Handbook 4th ed. New York: McGraw-Hill, 1998. Adapted with permission.

Figure 2-1. Sample form of the completed FMEA analysis; blood transfusion process (Burgmeier, 2002).

2.1.4 Miscellaneous Approaches to Risk Assessment

Outside the realm of informal and formal risk assessment regimes, which have been the trend since the publishing of the IOM report, literature also contains a number of retrospective studies on some adverse events, titled under the umbrella of “risk assessment”. The core of these studies is usually a linear regression between the adverse event and a few clinical factors. For example, Fortinsky et al. (2004); assess the risk of falls finding balance disturbance, multiple medications, sensory deficits, environmental hazards etc., among dominant influencing factors. Mrdovic, et al. (2011), use regression analysis to determine predictors of 30-day major adverse cardiovascular events (MACE) after primary percutaneous coronary intervention (PCI), and based on these factors propose a scoring system to assess the risk. Calvillo-King, et al. (2010), proposes a scoring system to predict probability of death or stroke after carotid endarterectomy in asymptomatic patients. Ammann, et al (2010) develop a scoring system to predict risk of adverse events (i.e. serious medical complications, infection, etc.) in pediatric patients with cancer who experience fever and neutropenia (FN). We couldn’t find any studies in which the investigators attempt to predict the risk of an adverse event. Instead, they tend to be retrospective/descriptive, deconstructing adverse events and trying to simply identify what factors might have contributed to their occurrence. The weakness in all of the published studies is, that there is not a systematic assessment of a control group. In other words, there is no effort to determine how frequently the so-called contributing factors were present in cases that did NOT result in an adverse event/outcome.

Also among these miscellaneous methods, are a few checklist types, and scoring system approaches that create simple numerical scoring systems that categorize patients by their susceptibility to certain adverse events. Many scoring systems to categorize patient's risk of developing pressure ulcers (or bedsores; an area of skin that breaks down due to constant pressure against skin) have been developed, which are discussed in more detail in section 5.1.1.1.

However, the reliability and validity of these scoring systems are not clear. For instance, in the case of risk scoring systems for pressure ulcer, some experts believe that often people who are identified as high risk for say pressure ulcer, do not experience pressure ulcer since resources would be dedicated to them to prevent the adverse event, but on the other hand patients with low scores in pressure ulcer risk end up developing pressure ulcers since some precautionary interventions might be ignored because they have been identified as a low risk patient.

2.2 Formal Risk Analysis Methods

Probabilistic Risk Analysis (PRA) is a systematic methodology to assess the risk of complex systems, and is currently being applied to many sectors from chemical processing to financial management. It has also had limited use and application in healthcare domain.

In many cases human performance, cognition and decision making are also involved in the performance of complex systems. Since humans can both initiate and mitigate the severity or the likelihood of accidents, the influence of humans on system risk and

reliability must be considered for a comprehensive PRA (Bedford and Cooke, 2001). Human Reliability Analysis (HRA) consists of a set of tools and methods that generally assess the probability of human error in certain tasks. Due to complexity and difficulty of quantifying human reliability, as compared with determining the reliability of mechanical or electrical components, extra steps are involved in modeling and quantifying the human element, before it can be used as an input to standard PRA tools such as fault trees or event sequence diagrams. A number of studies in healthcare domain are only focused on assessing the reliability of human elements (the HRA element of PRA). In our literature review, we have separated these studies from more systemic PRA studies in healthcare.

2.2.1 PRA in Engineering

Probabilistic risk analysis originated with the Reactor Safety Study WASH-1400 in the 1970's. PRA is mostly used in high-risk industries such as nuclear power plants, aviation and chemical industry. PRA provides a formal systematic way to identify and represent the factors that contribute and the chain of events leading to adverse events in complex technological systems. These factors include hardware failure, software failure, and human actions, interactions between involved parties and organizational factors (Wreathall and Nemeth, 2004, Stamatelatos, 2002). Common tools used in a conventional PRA are Event Trees (ET), Event Sequence Diagrams (ESD) and Fault Trees (FT).

The PRA ultimately presents a set of scenarios, frequencies and associated consequences. A scenario (represented by an ESD or ET) contains an initiating event (IE) and one or more pivotal events leading to an end state. Each pivotal event must be modeled in

sufficient detail to support valid quantification of scenarios. ESDs and ETs, use a forward reasoning logic that works forward through a causal path to model risk.

Complex pivotal events are frequently modeled using Fault Trees (FT). A FT is a picture of a set of logical relationships between more complex events such as system level failures, and more basic events such as component level failures. FTA, in contrast with ETs and ESDs, uses a backward reasoning, deductive and top-down logic, that deconstructs the top event (a failure) to the elements that cause and contribute to the occurrence of the top event. FT modeling is applicable to modeling hardware failure as well as other complex event types such as software failure and crew action (NASA PRA guide). Figure 2-2 depicts the typical format of a classical PRA methodology (Stamatelatos, 2002).

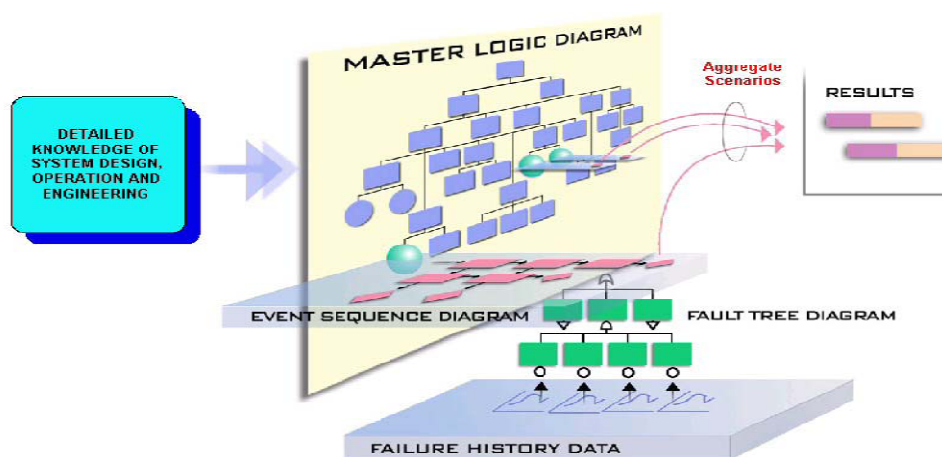


Figure 2-2. Classical PRA methodology (Figure originally composed by Futron corporation, NASA contractor for ISS PRA)

2.2.2 PRA in Healthcare

Although PRA has proven to be very effective in high-risk industries, mainly, nuclear power plants, a small number of studies using formal risk assessment tools (i.e. event trees, event sequence diagrams, fault trees) have been published in healthcare. That may also be because of the differences that exist between healthcare and the domains that PRA has traditionally been applied to. The diversity of the medical procedures, the treatments specific to a patient, the wide range of medical personnel are among the most notable of these differences. Exploring the use of PRA in anesthesia (Pate-Cornell, et.al. 1997), PRA in radiation brachytherapy (Ostrom, et.al., 1994), a fault tree analysis to understand why people deviate from prescribed protocols (Hyman, 2005), a fault tree to model risk in distributed healthcare information system (Maglogiannis and Zafiroopoulos, 2006) and a model for medication system failures in long-term care facilities using PRA (Comden, et.al., 2005) are among the studies on PRA application in healthcare.

2.2.3 Shortcomings of PRA in Medical Applications

Formal PRA modeling tools typically represent top-level failures or faults (termed adverse events in the medical domain) as the outcome of a linear sequence of events or component failures. This is by no means the case about the risk scenarios in healthcare. Much of what happens in healthcare is subject to feedback. For instance, an initiating event might not lead to an adverse event at time “t”, but because of the reinforcing effect of feedback might end up leading to an adverse event at time “t+n”. Additionally, the

number of contributing factors in the healthcare domain, are much greater than in mechanical systems. Even though conventional PRA methods can be useful in modeling specific aspects of healthcare-related risks, such as medication error which is a much more linear process, they are not adequate for modeling risk in healthcare for aforementioned reasons.

2.2.4 HRA in Engineering

Human Reliability Analysis can be considered as an extension of human-factors engineering that is basically concerned with identification and classification of human error and causalities involved and the prediction of operator performance. The common methods used are mainly cognitive control based techniques such as Contextual Control Model (COCOM) (Hollangel, 1993), Cognitive Reliability and Error Analysis Method (CREAM) (Hollangel, 1998). While PRA has been used in high-risk industries including nuclear power plants and aviation for the past thirty years to develop an understanding of risks involved in complex systems and the underlying causalities, HRA as an important part of the PRA, has increased the understanding of human performance issues that affect risk and safety in such systems.

2.2.5 HRA in Healthcare

There have been few if any well-designed efforts to understand human reliability and performance in healthcare settings. Instead, as evidenced in the published literature, investigators or theorists have mostly reviewed what historically was done in engineering domains, ‘cherry picked’ parts of existing theory or ‘cherry picked’ tools that might be

applicable to a small part of the immensely diverse set of human tasks and activities in healthcare. It is important to emphasize that:

- Healthcare is an enormously diverse domain
- The tasks are extraordinarily diverse, involving both skill-based and cognitive functions
- The humans performing these are diverse
- The ‘plant’ response, or ‘system’ response to a human’s actions are often unpredictable and subject to random as well as yet-to-be defined factors. This makes it really difficult to measure how much of the outcome was due to the human’s performance
- There is a wide range of tolerance to incomplete or imperfect task execution and it is really context sensitive – in other words, in some cases, the precision of a surgeon’s actions with a scalpel and suture may make the difference between life and death; other cases, it may not impact the overall outcome at all.

These are some of the reasons why there is not a unified set of theories or tools to confidently assess the human contribution to system safety

While human reliability analysis (HRA) has been well established and integrated into safety analysis in other industries (nuclear, aviation...), its application to healthcare is limited (Lyons, et.al. 2004). HRA studies human operator performance in the context of a specific task environment. It is often focused on estimating the probability of human

error, and how this probability might increase or decrease when coupled with various performance shaping factors.

Some work has been done trying to identify performance-shaping factors that are either unique or applicable to healthcare domain (Vincent 2000, Carthey et.al.2000). There have been a few foundational research activities directed at formally studying human reliability or performance shaping factors that might be unique to healthcare domain and healthcare transactions. For instance, Lyons, et al., 2004, conducted a literature review and lists 35 HRA primary techniques that might have a potential application in healthcare, based on common and general tasks in healthcare environment. They group these techniques into five categories of techniques for formally measuring performance in either controlled or naturalistic settings, as appears in Table 2-1. Note that these are not theories that might be useful in understanding performance shaping factors that contribute to human error.

Type of Technique	Description
Data Collection	Collection of information on incidents, goals, tasks, etc.
Task Description	Taking the data collected and portraying this in a useful form
Task Simulation	Simulating the task as described and changing aspects of it to identify problems
Human Error Identification and Analysis	Uses task description, simulation and/or contextual factors to identify the potential errors
Human Error Quantification	Estimates the probability of error identified

Table 2-1. Categories of HRA techniques (Lyons, 2004)

Data collection techniques used in HRA have also been used in some healthcare applications. What is important to understand though, is that some healthcare tasks, with well-defined bounds for correct and incorrect performance, such as pharmacy dispensing. Applying a relevant and useful technique for many other tasks in healthcare that are

messy, with poorly defined bounds for correct or incorrect performance, poorly defined end points, and are influenced by the feedback effects is extremely challenging. As an example, American Society of Health-System Pharmacists (ASHP) conducted a national survey of drug dispensing and administration practices (Pederson, et.al. 2002). An analysis of human error in anesthesia (Nyssen, 2001) and studies on incident reporting in anesthesiology (Staender et.al. 2001) and Case record review of adverse events (Woloshynowych, et.al.2003) are also among these studies.

Another study sets out to document the nature and incidence of surgical errors enacted during laparoscopic surgery using Systematic Human Error Reduction and Prediction Approach (SHERPA), which is a technique involving task analysis (Joice, et.al.1997). The same approach also has been used (Malik, et.al, 2003,) to detect surgical error in endoscopic DCR surgery.

Some of the more recent HRA studies in healthcare include; Inoue and Koizumi (2004), developed a model called EDIT (Error type, Direct threat, and Indirect threat) to characterize individual errors by evaluating error type, performance shaping factors (direct threats), and organizational factors (indirect threats) and applied this model to nursing practices in six hospitals. They find violation of rules, failure of labor management and defects in the standardization of nursing practices to be the three major organizational factors underlying medical error. Phipps et al. (2010), use a social psychological approach to investigate the anesthetists' beliefs about clinical practice guidelines to study determinants of intention to deviate from clinical practice guidelines. Gauba et al. (2008) and Cox et al. (2008) conduct studies to identify and quantify human errors in cataract surgery. Chadwick and Fallon (2011), use a modified version of Human

Error Assessment and Reduction technique (HEART) to analyze Record Abnormal Blood Results, a critical nursing task in radiotherapy treatment.

Other HRA techniques such as Cognitive Reliability and Error Analysis Method (CREAM), and Technique for Human Error Rate Prediction (THERP) have not been applied in healthcare domain.

2.2.6 Shortcomings of HRA in Healthcare

The main reason that human reliability analysis has not caught on in healthcare domain as well as it has in industry, is that healthcare is very different in some respects, despite some similarities) and it cannot be treated the same way as a nuclear power plant or a chemical plant. In a broader sense, power plants, aviation, chemical plants and healthcare are high-risk complex activities performed in large complex organizations, some aspects of healthcare are closer to some industries in comparison. For instance a pilot's work is similar to the high-tech monitoring of an anesthetist, but very different from what a surgeon does (Lyons et.al. 2004). There are also profound differences between healthcare and other high-risk industries. Healthcare consists of extraordinarily diverse set of activities.

Routine surgeries can sometimes be unpredictable and potentially harmful; treatment of acute psychosis may require quick decision making and response to the possible violent or bizarre behavior of the patient. Considering the wide range diversity of tasks in healthcare, some routine such as blood work, others as unpredictable as emergency medicine, and one can realize that the comparison with other high-risk industries with usually a limited set of activities is not a very meaningful comparison. Additionally, there is more uncertainty involved in healthcare practices than it is in industries such as nuclear

power plants where tasks are ideally, routine and deviation from usual practice is unusual and is to be avoided. For instance a patient's disease may be masked or difficult to diagnose or the result of the tests might not be clear. There is a higher level of uncertainty tolerance expected in this domain than other industries. Also most of interactions in healthcare are human-human as opposed to than human-machine interactions in other high risk industries.

Moreover, HRA focuses on the individual operator's performance in a controlled environment, and it is only as good as the level of expandability of this test environment to the real world setting. Conventional HRA methods also do not offer a causal picture of operator error. HRA approach provides a very limited insight when implemented in healthcare domain, since the results of analysis performed in controlled settings with individual operators and predefined tasks are only relevant to a very small portion of medical procedures.

3 Methodology

3.1 Overview

The approaches that have been adapted from engineering discipline and industry, reviewed in the background section, have had limited utility when used to model system-based risk in the healthcare domain. In particular classical PRA framework looks at risk scenarios as a linear chain of events that lead to an unsafe condition, which is by far not the case in healthcare. Most of the underlying causal chains in healthcare which result in an adverse event are subject to feedback and also the number of contributing factors is much greater than that of mechanical systems, and the magnitude of effect is non-linear.

Hence, In order to realistically model system-based risk in healthcare settings, it is necessary to account for dynamic factors and reinforcing loops, display the complexity of contributing factors, capture feedback and incorporate temporal factors.

The modeling approach adopted here consists of two components: a system dynamics framework and a Bayesian belief network (BBN) structure. This formalism has been introduced in Mohagheh, et.al. (2008) and applied in aviation safety context. The system dynamics formalism enables us to represent change over time and change due to feedback. The Bayesian belief network formalism enables us to represent networks of causality and capture stochastic characteristics of the system and the uncertainty related to that. BBNs also enable us to incorporate new knowledge and update the model as new evidence becomes available. The next sections briefly explain the components of the proposed model and the advantages that the combination of the two offers to more accurately and realistically capture risk dynamics in the healthcare domain.

3.2 Bayesian Belief Networks (BBN)

3.2.1 Introduction

Probabilistic networks in general, are graphical models that depict causal relations and interactions between a set of variables, where nodes in the graph represent variables and arcs or edges represent direct connections (direct dependencies) between the nodes Figure 3-1. If a pair of nodes is not connected, independence between the variables represented by these two nodes is assumed. Graphical models are intuitive and compact representations of (causal) dependencies and independencies between problem-domain variables. The advantage of graphs in probabilistic modeling is threefold; to provide convenient means of expressing modeling assumptions, to facilitate representation of joint probability functions and to facilitate efficient inferences from evidence and observation (Pearl 2009).

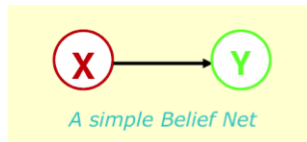


Figure 3-1. A simple belief network; event X (parent node, cause node), causes/influences event Y (child node, effect node)

More specifically, Bayesian Belief Networks or Bayesian Networks are a class of probabilistic graphical models for reasoning under uncertainty, where nodes represent discrete or continuous variables and arcs represent direct causal connections (relationship) between them. The graphical aspect of probabilistic networks can be used in a qualitative manner to represent relationships between a set of variables (Kjaerulff

and Madsen, 2008). While the arrangement of the nodes and arcs of the graph/network structure can represent the qualitative relationships between variables, the *strength* of the causal relationship between variables on the other hand, can also be quantified using probability calculus. Specifically, each variable (node) in the network is represented as a finite set of mutually exclusive states, and a conditional probability table can be created for each variable (node) and its parent(s), by the conditional probability distributions associated with each node. This probabilistic and numerical aspect of probabilistic networks is referred to as *quantitative* aspect. To elaborate, in the simple Bayesian net in Figure 3.2, nodes X and Y are called parent nodes or input nodes and node Z is called the child node or the target node. If we assume binary states for each of these nodes;

$$X = \{x, \bar{x}\}, Y = \{y, \bar{y}\}, Z = \{z, \bar{z}\}$$

with probability distributions

$$P(x) = q_x, P(y) = q_y, P(z) = q_z,$$

our objective is to calculate q_z as a function of q_x and q_y .

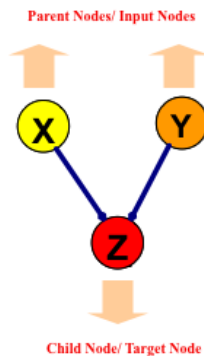


Figure 3-2. Input and output nodes

The only constraint is that BBNs are directed acyclic graphs (DAG), meaning that starting from a node, you cannot return to that node simply by following the directed arcs.

BBNs, which are compact networks of probabilities that capture probabilistic relations between variables and contain historical information about their relationship, have proven to be powerful tools for modeling causes and effect in many domains. They are also very effective in modeling situations where data are uncertain and vague or incomplete and only partially available. This uncertainty in information can arise in many situations; domain experts may be uncertain about their knowledge, there might be uncertainty about the accuracy and/or availability of the information or the situation being modeled might be inherently uncertain. (CRA, 2004)

3.2.2 BBN Elements

To explain the structure of BBNs and how they are built and used for inferences, we will use a previously published example of a Bayesian net intended to support a medical diagnostic task for lung cancer.

The example, through which we will review the structure of BBNs in this section, is a simplified and modified version of a problem known as Asia problem (Lauritzen and Spiegelhalter, 1988, and Korb and Nicholson 2004).

A patient, who is experiencing shortness of breath, visits his doctor in fear of lung cancer. The doctor knows that possible candidate diseases that may cause shortness of breath or dyspnoea, are lung cancer, tuberculosis or bronchitis.

Dyspnoea can also be due to presence of a number of these candidates or none of them. She also knows that smoking is a risk factor for both lung cancer and bronchitis, and that exposure to air pollution can be a contributing factor in lung cancer. The positive result of a chest X-ray would indicate either tuberculosis or lung cancer. The doctor would like to know the chance that lung cancer is present.

3.2.2.1 Structure of BBNs

The structure of the network represents the qualitative relationships between different variables. Two variables (nodes) are connected if one affects or causes the other and the connecting arc indicates the direction of the effect. For instance in our simplified example, we have assumed factors that affect a patient's chance of having lung cancer, are pollution and smoking. Similarly having lung cancer will cause breathing problems and will increase chances of a positive X-ray result. Hence the list of variables, their types and a set of possible values or states (chosen arbitrarily for this example) and the structure of the network will be as appears in Table 3-1 and Figure 3-2, respectively.

Node	States
Pollution (P)	Low, High
Smoker (S)	True, False
Cancer (C)	True, False
Dyspnoea (D)	True, False
X-Ray (X)	Positive, Negative

Table 3-1. List of variables and their states

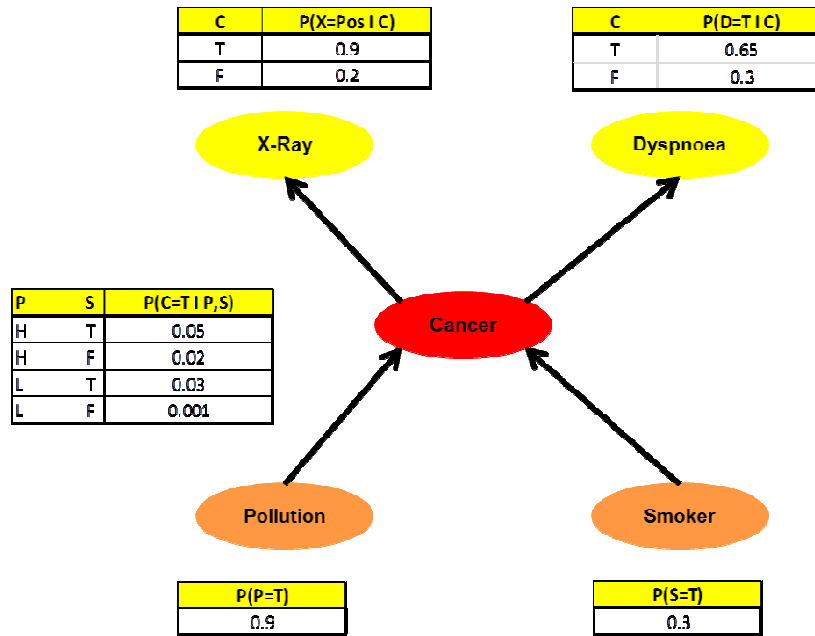


Figure 3-2.Lung cancer BBN

3.2.2.2 Conditional Probabilities

Bayesian networks have a qualitative aspect and a corresponding quantitative aspect. Once the qualitative model (graphical representation) is established, we will need to quantify the strength of the relationship between connected nodes, by assigning a conditional probability table, CPT, (in form of a distribution for continuous variables or a point estimate for discrete ones) to each node. Conditional probabilities represent the likelihood based on historical data or our prior knowledge and belief.

Mathematical representation of conditional probability is $P(x|p_1, p_2, \dots, p_n)$, that is the probability of variable X being in state x , given the states of its parent nodes P_1, P_2, P_3 are in the states p_1, p_2, \dots, p_n , respectively. Therefore for each parent and each possible state of that parent, there exists a node in the conditional probability table that indicates the

likelihood that the child will be in some state. In our lung cancer example, from Figure 3-2, we can read $P(\text{Cancer} = T | \text{Pollution} = H, \text{Smoker} = T) = 0.05$.

3.2.2.3 Inference with Bayesian Networks

One of the most important features of Bayesian networks, is that can be used for updating our prior beliefs and calculating new beliefs as new information and observations or in other words “*evidence*” becomes available. In fact, up until this point, there is nothing Bayesian about the Bayesian networks. As Bayesian networks often represent causal relationships of the nature of $X \longrightarrow Y$, the task of inference is then to derive the posterior probability distribution of Y , $P(Y|X)$, where X is a set of observations (evidence), and Y is the variable that is important for prediction or diagnosis (Pearl, 2009). This is a straightforward application of Bayes’s rule which yields:

$$P(Y|X) = \frac{P(Y = y|X)P(X)}{P(Y = y)}$$

Where $P(X)$, is our prior belief about X and

$$P(Y = y) = \sum_x P(Y = y|X = x)P(X)$$

For instance if the doctor receives a piece of information indicating that the patient has been exposed to high level of pollution, then this evidence is set in the Bayesian network as:

$P(P = H) = 1$, and consequently $P(P = L) = 0$. Using Bayes's rule from equation above, probability of the patient having lung cancer $P(Cancer = T | P, S)$ will increase from 0.02 to 0.03.

3.2.2.4 Types of Evidence

Evidence is any type of information about the current situation of a variable /node. For instance, in our example, if we find out that the patient is a smoker; our belief about the probability of him having a lung cancer will change. In general there are two types of evidence available for BBNs:

- Hard evidence: Assigns a zero probability to all but one state of the variable
- Soft Evidence: Bayesian networks also support evidence that is vague or incomplete or uncertain. This type of evidence is called soft evidence, which is any evidence that is not hard evidence. In our example, if the doctor knows that there is 90% chance that the patient has been exposed to pollution (but he is not 100% sure), in lung cancer BBN model he will assign a probability of 0.9 for the node "pollution" as evidence. If he knew for sure that the patient was definitely exposed to pollution (i.e. had hard evidence) he could assign a probability of 1 to this node.

3.2.3 Construction and Quantification of BBNs

A Bayesian belief network could be constructed manually (i.e. Based on expert knowledge, literature, etc.), automatically from data (i.e. Data driven BBN construction) or through a combination of manual and data driven approaches. To induce the structure of the network i.e., the graph, from a source of data there exist different classes of algorithms such as search and score algorithms, constraint based algorithms and combinations of the two (Kjaerulff and Madsen 2008), which require considerable amount of data. The BBNs constructed in this thesis are built using manual approaches.

When faced with a problem, the first step would naturally be figuring out whether Bayesian networks are the right tool and approach for the problem based on the nature of the problem. Generally, when dealing with problems where there is uncertainty associated with the cause and effect relations and mechanisms; Bayesian belief networks seem to be the ideal framework.

A Bayesian belief network has two major components; the structure and the parameters (i.e. conditional probabilities). The structure of a BBN is usually referred to as the qualitative part whereas the parameters and the conditional probabilities are the quantitative part. Consequently the model elicitation process consists of two phases; first the variables and the causal relations between them are identified and second, once the structure of the model (i.e. the graph) has been established and verified the values of the parameters and conditional probabilities are elicited. The manual construction of the Bayesian net could be a labor-intensive task requiring some level of creativity and also close interaction with domain experts.

3.2.3.1 Construction of the BBNs

The qualitative part of building a Bayesian network involves identifying the variables (i.e. nodes in the graph) and identifying the causal relations between the variables (i.e. the edges or the arrows). Kjaerulff and Madsen (2008) categorize variables types into four different classes.

- Background variables: usually the root variables of a Bayesian network
- Problem variables: the variables of interest, for which we want to compute the posterior probability distribution given the observations
- Mediating variables: directly unobservable variables for which posterior probability is not of interest but they play an important role in establishing accurate conditional independence and dependence relations in the model and are most often influenced by the background and problem variables
- Symptom variable: observable as the consequence of the presence of problem variable and influenced by it

Given the above classification, typically the overall causal structure of a Bayesian network will be as depicted in Figure 3-3.

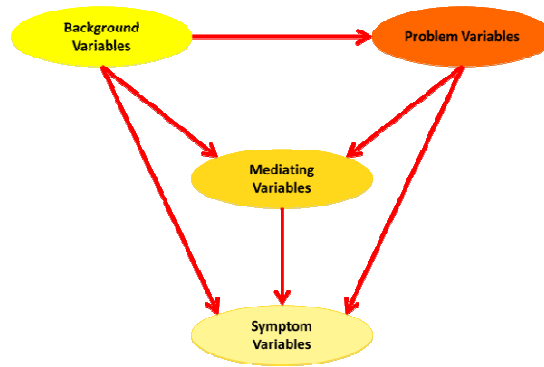


Figure 3-3. Overall causal structure of BBN (Kjaerulff and Madsen, 2008)

Kjaerulff and Madsen (2008) also present two structured ways of eliciting model structure in their book. First, a basic approach that relies on the causal ordering that exist between the variables and identifying model variables (in any of the above mentioned categories) and identifying the causal links between these variable. Second, a more refined approach that constructs models using five commonly occurring substructures called idioms. A vast majority of Bayesian nets are claimed to be constructible using these idioms or substructures. Our approach in this thesis to build the Bayesian networks is the basic approach, hence the readers interested in the second, more refined approach referred to Kjaerulff and Madsen (2008) and Neil et al. (2000).

One should note that building a network often requires a careful and delicate trade-off between desire to build a large, rich and super comprehensive model that covers every little detail to obtain the highest level of accuracy possible on one hand, and the feasibility and the cost of construction, and the complexity of probabilistic inference on the other hand (Druzdel and van der Gaag, 1995).

3.2.3.2 Quantification of the BBNs; Eliciting the Numbers

After establishing the structure of the network (the qualitative part) through the iterative process of model verification revision, identification of new variables, deletion or modification of the existing variables, and the addition or deletion of the causal links (edges), the next and probably the most challenging phase is the elicitation of the conditional probability distributions or populating the conditional probability tables (CPTs). The amount of effort that goes into building the structure of the model and even more so into obtaining the numerical parameters, is probably the biggest obstacle in the way of applying Bayesian nets in many practical problems (Onisko, Druzdzel and Wasyluk, 2001).

Since the process of eliciting the quantitative information required for this stage of constructing BBNs is often very demanding, it is important to carefully verify the structure of the Bayesian net before proceeding to the quantification phase. Nevertheless making minor adjustments to the structure of the network, in order to reduce the number of parameters, is sometimes inevitable. The parameters of the network can be obtained from databases in literature or elicited from subject matter experts (Druzdzel and van der Gaag, 2000).

For the variables that field data exists, the task of computing the marginal and conditional probabilities is quite straightforward. Very often there is incomplete or no data available and the analyst has to rely on the subjective assessment of probability obtained from domain experts (Diez, 93).

In this section we will focus on the most intimidating task in building Bayesian networks, obtaining the required probabilities.

3.2.3.2.1 Sources of Information

In most applications, probabilistic information is available through one or more of three sources; Statistical data (field data), Literature and Subject matter experts. In data rich applications it is usually not too difficult to collect data, on the variables of interest. If comprehensive data is available, both the qualitative part (the graph) and quantitative part (the probabilities) can be automatically constructed. There are two approaches to learning the structure (graph) of the Bayesian network from data. First the constraint-based search and second Bayesian search for graphs with highest posterior probability given data. Since we have not constructed our BBNs learning from data, we will not cover these approaches but for more information on the former please see Pearl and Verma, 1991, Spirtes et al., 1993 and for the latter please see Cooper and Herskovits, 1992. Once the structure of the network is established, the task of acquiring probabilities will consist of studying the subsets of data that correspond to the various conditions (combinations of various states of the variables).

However, in most cases where reliable statistical data is scarce other forms of data should be considered. Literature often provides a good source of probabilistic information. For instance, more specifically in the field of medicine, many studies report on the disorders and symptoms and the causal relations between them, but one has to be careful since this probabilistic information are not always directly useable in Bayesian nets. For example, one could find the conditional probabilities of symptoms given the disorder or the disease

is present, but conditional probabilities of symptoms given the disorder is absent are rarely reported. Also most often the probabilities required for the intermediate disorder states that have been modeled in the network are not studied or reported (Druzdel and van der Gaag, 2000 and 1995).

Finally, if there are few or no reliable data available experts' knowledge and experience is used as a source of probabilistic information. Although the role of experts in providing the parameters of Bayesian nets and the probabilities should not be underestimated, the problems and challenges in eliciting probabilities from experts which have been discussed in many books and articles should be acknowledged. Nevertheless, many techniques have been developed for eliciting well-calibrated, unbiased and reliable probabilities from domain experts (Druzdel and van der Gaag, 2000, O'Hagan, et al., 2006).

Bayesian networks typically consist of tens and sometimes hundreds of variables (nodes) and hence require hundreds of probabilities, and a good part of these probabilities –if not the majority- has to be assessed by domain experts. Given that expert's time is an expensive commodity supplementary techniques have to be utilized to reduce the burden on the experts.

The amount of information and the number of probabilities to be elicited is dependent on the structure of the graph and the number of variables in the graph. The number of required probabilities grows exponentially with the size of variable's parental set. To reduce the number of probabilities to be elicited, two approaches are commonly used. The first approach is based on the modifications made in the structure of the Bayesian

net, i.e. the graph, and the second approach is based on using parametric probability distributions. The first approach uses techniques such as parent divorcing and introducing an intermediate variable, temporal transformation, etc. to adjust the structure of the model with the goal of easing the process of eliciting the probabilities and making network quantifications manageable. These techniques have been discussed individually or collectively in Kjaerulff and Madsen (2008), Olsean et al. (1989), van Engelen (1997) among others.

Noisy-OR gates, Noisy-AND gates and their generalizations (Drudzel and van der Gaag, 2000, Heckerman and Breese, 1996, Pearl, 1988, Lemmer and Gossink, 2004) on the other hand, are examples of using parametric probability distributions to reduce the number of probabilities to be elicited. Methods based on this second approach, are based on the assumption that the parents of a variable in the network are causally independent. With these methods, the number of probabilities to be assessed for a variable grows linearly rather than exponentially as the number of its parents increase. For instance, in Noisy-OR gate, for a node that has “n” parents with binary states, the number of probabilities to be elicited is “n” rather than “ 2^n ”. That is for a node with 10 parents, we only need to ask experts for 10 probabilities rather than 1024 probabilities, using the Noisy-OR model.

Using these two approaches, modifications to the structure and/or parametric probability distributions, however, will probably compromise the accuracy of the model but then again as mentioned before we are dealing with a trade-off between accuracy and feasibility (by carefully reducing the model using the above approaches) in building Bayesian networks.

Although the process of eliciting the probabilities is done after the analyst reaches a steady, reliable and robust structure for the Bayesian net, to think that this elicitation is a one-shot process is rather unrealistic. That being said, literature could be found on the ways and tools to support the elicitation process. For instance, after collecting the first round of probabilities, which is probably raw and less accurate and less calibrated, using the sensitivity analysis the analyst would be able to discover the most important probabilities and refine them (Coupe et al., 2000, Philips, 1982). Also this process highlights the less influential probabilities in the network that could probably be eliminated or further simplified without seriously compromising the accuracy of the model. This process is done iteratively until the cost of further refining elicitation outweighs the benefits of more accuracy achieved, or the till the accuracy could not be improved any further simply because we don't have further knowledge available.

Given the scarcity and the high value of expert time, this will help focus the efforts and resources on the parts of the model that simply put, matter most.

3.2.3.3 Construction and Quantification of the BBNs in This Research

To develop the BBNs in this research we first started with a set of factors in the literature and one of the experts added, deleted and modified these factors and identified the causal relations between the factors, which resulted in the first draft of the BBN. This first draft was then discussed with domain experts using the interview guides in Appendix A and Appendix B in multiple sessions and each expert provided their opinion about:

- The variables of the model; whether they thought there should be other factors considered. If they had other factors they added to the model and made any modifications to the existing factors they thought were necessary.
- The causal relations between the variables

After incorporating all the changes made to the model by the experts the latest version of the model was discussed with the experts in another interview and the experts were asked to score the model in the scale of 1 to 100 in the terms of model completeness, model accuracy, ease of understanding and perceived predictive validity, to ensure sufficient confidence in the structure of the model before proceeding to model quantification. We have discussed this further, in BBN validation and verification part of this dissertation, sections 5.1.4.2, and 5.2.4.2.

In quantification of the Bayesian Belief Networks in this thesis we have used both structural techniques and parametric probability distribution techniques, which have been explained in detail in section 4.1.3.1.

3.3 System Dynamics (SD)

3.3.1 Introduction

A rather popular approach to understanding the behavior of complex social and economic systems is the application of non-linear differential equations (DEs). System dynamics (SD) is a simulation based, differential equation modeling tool that is widely used in situations where the formal model is complex and an analytical solution is impossible or very difficult to obtain (Sterman, 2000). It is a method to enhance learning in complex systems. Just as airlines use flight simulators to train their pilots, SD develops “management flight simulators”, to help us learn about dynamic complexity, predicting the impact of policies and decisions, understand the sources of policy resistance and design more effective policies. System dynamics is fundamentally interdisciplinary. Since our concern is the behavior of complex systems, SD has its roots in the theory of non-linear dynamics and feedback control developed in mathematics and engineering. Because we apply these tools to human behavior as well as technical systems, SD also draws on cognitive and social psychology, economics and social sciences as well, that helps us better understand the sometimes counterintuitive behavior of social systems (Sterman 2000, Forrester1975). The purpose of building system dynamics models is to explain and understand the behavior of complex systems and how they evolve overtime, since they can take into account (multiple) feedback mechanisms and non-linear relationships between system variables.

Over the years, SD has been applied to a wide variety of situations, ranging from corporate strategy to the dynamics of diabetes and from cold war arms race to HIV

combat with human immune system. It can be applied to any dynamic system, with any time and special scale. In the case of our research problem, for capturing the effects of organizational decisions and policies on risk of adverse events, the feedback effects and nonlinearities, system dynamics formalism is a well suited and efficient tool.

3.3.2 Building Blocks

There are two major building blocks of system dynamics models: stock and flow diagrams and feedback or causal loops. Below, we will provide a brief overview on these building blocks which will hopefully facilitate the interpretation of the system dynamics model proposed in this research.

3.3.2.1 Feedback loops

Feedback is one of the core concepts of system dynamics and our mental models often fail to include critical feedbacks that determine the dynamics of our systems. These feedbacks are modeled using causal loop diagramming in system dynamics.

Feedback processes take place, if a system component (variable) initiates changes in other components (variables) of the system that in return; affect the very component that had originally initiated the change (Ruth, 2001). This process usually occurs through non-linear relations between system components and can involve time delays.

An essential part to system dynamics modeling is to understand and present the feedback mechanisms, that along with the stock and flow structures, nonlinearities and time delays

form the dynamics of a system. Most complex system behaviors are usually due to the feedback relations between the components of the system rather than the complexity of the components themselves. (Sterman, 2000)

There are only two types of feedback that form all dynamics; Positive (self-reinforcing) feedback and negative (self-correcting) feedback. More complex interactions may be captured through a combination of these two types. Positive loops, cause reinforcement or amplification of the events in the system. For example, the more money you invest in a savings account, the higher interest you will receive.

Negative loops, on the other hand, oppose the change. Negative feedback processes usually lead systems towards equilibrium states (Ruth, 2001). The less the strength of a pain killer, the more pills you have to take to soothe your headache.

Figure 3-4, is a very simple illustration of these loops. More eggs result in more chickens (figure 3 A, reinforcing loop). The more the chickens cross a road, the higher the chances of them getting hit by cars, hence, the higher rate of mortality and fewer chickens (figure 3-5 B, balancing loop). Figure 3-5 C, shows a combination of reinforcing and balancing loops.

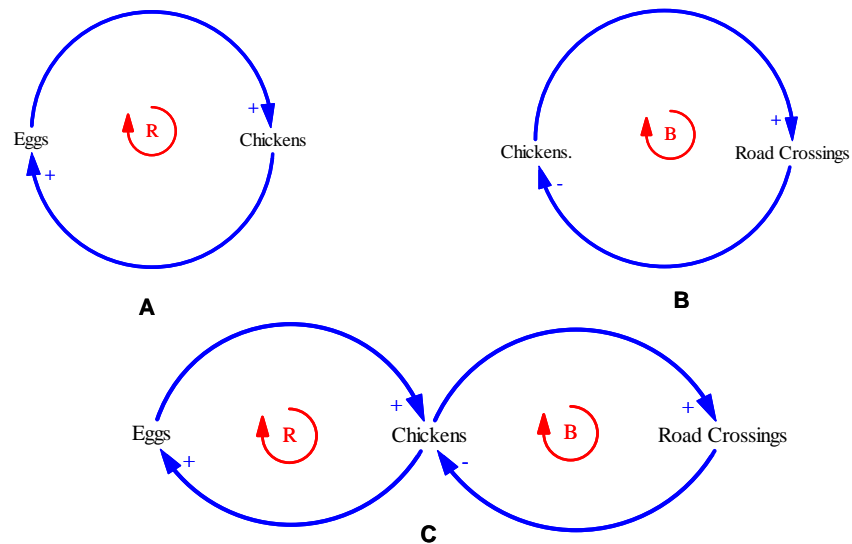


Figure 3-4. Balancing and reinforcing loops

Even though feedback loops are principally limited to two types, positive and negative loops, models can contain thousands of these loops interacting with one another with time delays and through nonlinear relations. The dynamic of the systems are the product of these interactions.

3.3.2.2 Stock and Flow Structure

Besides feedback loops, stock and flow structure is the other building block of any system dynamics model. Stock (population Figure 3-5) represents accumulation of some measurable entity (e.g. people, money, inventories of products or even intangibles such as happiness) (Ford, 1999). Stocks characterize the state of the system and generate

information upon which decisions are made (Sterman, 2000). Stocks change with the inflows and outflows.

Flows (Birth and death in Figure 3-5) are the physical or conceptual entities that enter or leave the system and move over time. Auxiliary variables (death and birth rate in Figure 3-5) help describe the flow.

Figure 3-5, represents a very stock and flow structure.

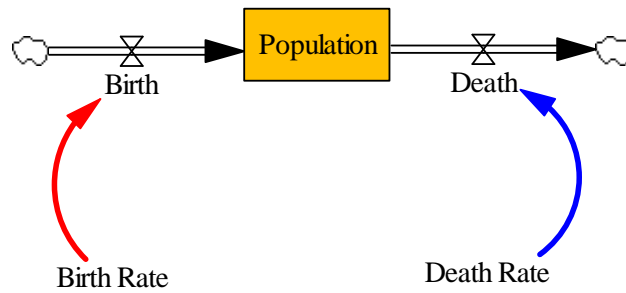


Figure 3-5. Stock and flow diagram

3.4 SD/BBN Combination

As mentioned briefly above, the system dynamics part of the model demonstrates how organizational level and policy level contributions to risk evolve over time, and how policies and decisions may affect the general system-level contribution to adverse event risk. Also, it captures the feedback of organizational factors and decisions over time and the non-linearities in these feedback effects. BBN part of the model, represents patient-level factors and also physician level decisions and factors in the management of an individual patient, which contribute to the risk of an adverse event.

Each patient, considering his or her individual medical condition and physician's decisions in treating this patient is exposed to a certain level of risk of specific adverse events (e.g. infection, pressure ulcer). This is captured with a BBN. On the other hand, the system dynamics section in the model, based on the financial situation of the hospital, level of dedication to safety and organizational and policy level factors and decisions with regards to staffing, pressure to reduce length of stay and investment in safety, which evolve dynamically over time, provides a background that determines where hospital is standing in terms of risk when the next patient walks in.

In our methodology, the system dynamics module (representing system level factors and decisions) and the Bayesian network module (representing patient level and patient-provider level factors), are integrated in a way that each module can provide input to a node(s) and/or receive input for its node from the other module allowing the entire hybrid environment to capture feedback and delay effects. The interface of SD and BBN can be captured by importing and exporting data from and to the system dynamics model. For

instance, the variable “Staff Adequacy”, that is an outcome of managerial decisions to reduce operational costs (captured in the system dynamics module), is also an important factor that may determine whether or not a patient is moved frequently enough by the staff which ultimately impacts patient’s “Risk of Pressure Ulcer” (captured in pressure ulcer BBN). So the input from system dynamics model to the Bayesian network for pressure ulcer is “Staff Adequacy”, and in return pressure ulcer BBN provides the updated value for “Risk of Pressure Ulcer” as an input to the system dynamics module.

Figure 3-6, depicts this interaction.

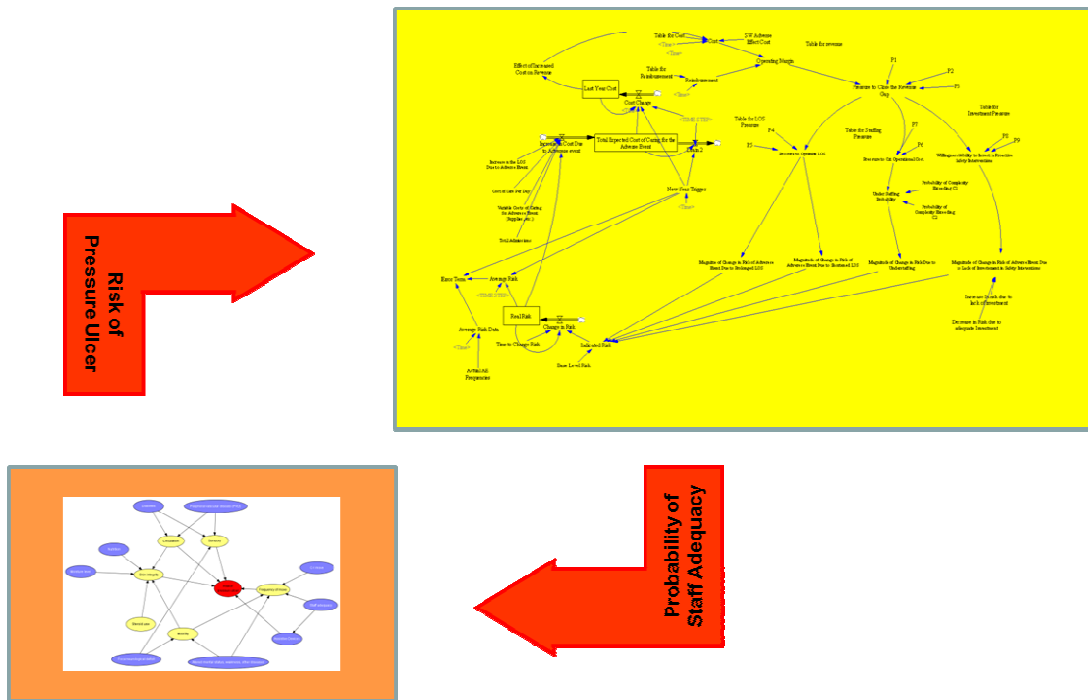


Figure 3-6. SD and BBN hybrid modeling environment

3.5 Information Sources for Quantification of Models

3.5.1 Data from Actual Operating Experience

Eight years of clinical data from one of Harvard Medical School's major teaching hospitals was made available for this study. Also will use data obtained from domain experts in all the steps of model development, quantification and validation.

Quantitative modeling will be informed by data stored in the administrative and clinical databases from a major teaching affiliate of Harvard Medical School. The clinical information system in the medical center has SQL servers that support 62 linked relational databases storing contemporary and historical clinical data (FY'99-FY'09) and disease registries for all major clinical areas (e.g., ED, Inpatient floors, outpatient areas, procedural suites, laboratory, pharmacy, radiology, etc.) as well as the more recently implemented computerized order entry data. The SQL Servers are accessible using the Microsoft Management Console tool kit and SQL Server Enterprise Manager Software. A series of SQL queries and stored procedures have been developed to extract the following data from these administrative data sources: patient and provider scheduling data for procedural and inpatient units; acuity levels and patients volume in both the target population as well as concurrent levels in other units of the hospital; total resource utilization at the unit and case level; nursing scheduling cycles; drug utilization (both standard and emergency pharmaceutical agents); laboratory results matched to the pre-, intra-, and post-procedural phases of care in procedural areas and; sub-process time stamps for procedural areas. The process data in these sources is remarkably detailed and will enable us to model durations of key phases of care such as

pre-procedure preparation, sedation phase, prep/drape phase, post-procedure recovery phase, admission and discharge times and room turnover times. Raw data for process durations, emergency case interruptions to the elective scheduled, delays in scheduled cases due to emergency issues, transition times between pre-procedure, procedure and recovery phases for each of the procedural units in the medical center, delay times and reasons for delays in initiating emergency interventions in these units as a function of day of the week and hour of the day. length of stay in recovery units as a function of time of day or proximity to shift change that are acquired using the stored procedures will be fitted to standard distributions using a Kolmogorov-Smirnov algorithm. The administrative data sources have been used extensively by risk analysts at the medical center to identify unreported adverse events or near misses using clusters of data as triggers. Examples of triggers include the identification of computerized order entries for blood product use in ‘low-bleeding-risk’ procedures, sedative reversal agents (e.g., flumazenil or naloxone used during the recovery phase) or physical restraints (suggesting agitation) in combination with specific procedures. We will use these cluster-based triggers to identify unreported adverse events, and update frequency estimates established from the self-reported events.

3.5.2 Adverse Event Data

Additional data is derived from an adverse event reporting system that currently contains approximately 10,000 cases, and 400 root-cause analysis reports that contain reconstructed causal sequences that will help inform the qualitative modeling phase of this work. All clinical data used in this study, has been de-identified and provided to us by one of the members of the advisory committee, and no direct

access to databases was possible for the author of this thesis due to confidentiality concerns.

The root cause analyses used at this medical center provide a rich classification of system and human factors thought to have contributed to the initiating or propagation of the event. It is unclear whether the frequency data are reliable to use for the quantitative analysis, since reported events do not accurately reflect prevalence or frequency. However, they serve as a fairly comprehensive source of data for the qualitative modeling of event initiation and propagation.

The adverse events that are of interest in this study are a set of twelve hospital acquired conditions (HAC) that patients could experience while in the hospital, which are thought to be preventable and Medicare is considering not to reimburse. Table 3-2 shows the list of these adverse events.

Selected HAC
Foreign Object Retained After Surgery
Air Embolism
Blood Incompatibility
Pressure Ulcer Stages III and IV
Falls and Trauma: (Fracture, Dislocation, Intracranial Injury, Crushing Injury, Burn, Electric Shock)
Catheter-Associated Urinary Tract Infection (UTI)
Vascular Catheter-Associated Infection
Manifestations of Poor Glycemic Control
Surgical Site Infection, Mediastinitis, Following Coronary Artery Bypass Graft (CABG)
Surgical Site Infection Following Certain Orthopedic Procedures
Surgical Site Infection Following Bariatric Surgery for Obesity
Deep Vein Thrombosis and Pulmonary Embolism Following Certain Orthopedic Procedures

Table 3-2. Hospital acquired adverse events

3.5.3 Expert Elicitation

Graphical tools could be used to support the network quantification process and eliciting the probabilities (Wang and Druzdzel, 2000). Graphical tools provide an interactive way to elicitation of probabilities. Allowing the expert to manipulate a chart, or choose from a set of functions that have been graphically presented will offer more support to the expert and is most likely help the expert to provide his estimate more confidently and more accurately. Probabilities could also be expressed through verbal expressions such as more very likely, certainly, or improbable (Renooij and Wittman, 1999), though verbal expressions could cover a wide range of numbers. Perhaps a combination of verbal

expression approach and the number approach could produce better results (Van der Gaag, et al., 1995).

Decision makers have always been interested in subjective knowledge using experts for their opinion. In fact, in many cases subjective knowledge may be the only source of information that exists for a particular problem of interest. The task of expert elicitation is ultimately threefold; selection of experts, elicitation of their opinions and judgments, and the aggregation of their opinions in the case of multiple experts. These tasks have been extensively discussed in the literature (Ayyub, 2001, O'Hagan, et al., 2006), but here we will briefly overview the process, with an emphasis on the needs of this research.

3.5.4 Who is an Expert?

When major decisions are to be made in presence of uncertainty and expert judgment is essential to minimize and characterize uncertainty, the choice of experts becomes one of the most phases of the elicitation process, and the success and usefulness of such process is directly dependent on the experience, knowledge and technical background of individual expert (O'Hagan, et al., 2006). Generally an expert could be defined as a skillful person with great knowledge of and extensive training in a specific field. However, to be precise, in the realm of expertise, there are many other psychological factors that may be determinant parameters in how a person uses and organizes his or her own knowledge. Interested readers are encouraged to see O'Hagan et al. (2006) and Wood and Ford (1993), among others, for detailed discussions.

Expert's opinion is then defined as expert's formal judgment on the specific subject within his realm of expertise. On the other hand, an opinion is a judgment, belief or

subjective assessment of the quality or quantity of the unknown of interest, based on uncertain information (Ayyub, 2001).

In selecting the experts it is important to understand that dependencies may exist between the experts. The analyst may try to recruit multiple experts from different organizations and backgrounds to reduce these dependencies, and eliminate the sources of strong dependencies. Weak dependencies however, do not seem to effect the value of expert judgment.

As a general guideline, Cooke (1991) formulates principals that should be considered in order for results to be considered scientific;

- Scrutability/accountability: All data, including experts' information and assessments, should be open to peer review and results must be reproducible
- Empirical control: Quantitative expert judgments should be subjected to empirical quality control
- Neutrality: The method for evaluating and aggregating expert assessments should encourage experts to provide their true opinion
- Fairness: Experts are not prejudged

In evaluating and processing experts' inputs in this project, we have been committed to these guidelines.

3.5.5 Elicitation Methods

When planning an elicitation, two issues need to be considered first. First, when we are seeking the subjective knowledge of an expert about an uncertain parameter, we would often like to gather opinions from several experts and consolidate their input into one probability distribution. If we collect experts' judgments separately we would need to use some type of algorithm to combine their opinions, and the process is known as *mathematical aggregation*. If we bring the group of experts together in the group and elicit one judgment from the group, the process is known as behavioral aggregation. The second issue is to decide whether the elicitation will be done through a face-to-face interview with the experts, or through using questionnaire. Face to face interview, should the means exist for the analyst, is without the doubt the best approach, since the analyst is present and could clarify any ambiguities and would be otherwise much more time consuming using a questionnaire. Also the analyst could much more effectively explain the model and the parameters to be elicited, and the type of expert input that would be most useful to the analyst. However, arranging interviews, especially individually with each expert, could be very challenging given that experts are usually busy professionals.

Ayyub (2001) attributes the first structured methods for expert opinion elicitation to the RAND Corporation in early 1950s. These two methods are Delphi method and Scenario analysis method.

Delphi method is probably the most known method of expert elicitation, which was developed for U.S. Air Force and used throughout the 1960s and 1970s in variety of

applications from technology forecasting and policy making to space progress and weapons systems. The basic Delphi method has 8 steps, which basically includes:

- 1- Developing questionnaires
- 2- Selection of experts
- 3- Familiarization of the experts with the issue of interest
- 4- Elicitation of experts opinion on the issue
- 5- Aggregation of experts' opinion
- 6- Review of the aggregation results by experts and revision of their initial opinion
- 7- Revision and review to achieve a complete consensus
- 8- Reporting results with justifications on the out of range opinions

Many elements and factors have been suggested in the literature as being critical to a good elicitation process (Clemen and Reilly, 2001, Walls and Quigley, 2001, Grathwaite, et al., 2005, NUREG 1150, 1989, O'Hagan, et al., 2006), but the heart and soul of all of these methods/processes are really the same. The steps that all have they common are; Preparation, Expert selection, training of the experts, Elicitation using appropriate format, Aggregation of experts' input. In the expert elicitation process we have carefully considered and used the processes suggested in the literature.

3.5.5.1 Aggregation of Experts' Opinions

A large number of methods exist for combining experts' opinion, from older methods like the Delphi method, to more involved Bayesian models. Among many Bayesian methods proposed in

the literature a few have had actual applications and even fewer have been applied more than once with the exception of proposed model by Mosleh and Apostolakis (1986) (Bedford and Cooke, 2001). The objective is to aggregate experts' point assessments for the unknown of interest, X. If x_1, \dots, x_e are estimates of X obtained from experts 1, ..., e, and our prior belief about X is expressed with $\pi_0(x)$, the updated belief about X, given the estimates x_1, \dots, x_e from the experts using Bayes' theorem is:

$$\pi(x|x_1, \dots, x_n) \propto L(x_1, \dots, x_n|x)\pi_0(x)$$

And assuming that experts are independent we could write the likelihood term as:

$$L(x_1, \dots, x_e|x) = \prod L(x_i|x)$$

The objective then reduces to determining $L(x_i|x)$. Mosleh and Apostolakis (1986) suggest to error models;

- Additive error model , $x_i = x + \varepsilon_i$
- Multiplicative error model , $x_i = x\varepsilon_i$

Where x_i is expert input, x is the true value of the unknown of interest X, and ε_i is the error term. The model also assumes that the error term has a normal distribution with mean μ_i , and standard deviation σ_i the decision maker has to choose these parameters based on each expert's bias and accuracy. Of course where past performance data is available, choosing these parameters is a much more straightforward task. Given these assumptions, the likelihood $L(x_i|x)$ of getting

estimate x_i from expert I, given that the true value of X is x, is obtained from a normal distribution with mean $x + \mu_i$ and standard deviation σ_i .

3.5.6 The Panel of the Experts

The experts were selected on the basis of their recognized expertise and experience in the field of medical risk assessment, patient safety, quality of medical care and also the specific adverse events that are of interest in this project. The experts were selected from a number of extremely reputable medical institutions including but not limited to, Harvard Medical School, Beth Israel Deaconess Medical Center, Johns Hopkins University School of Medicine, George Washington University School of Medicine, Federal Food and Drug Administration and Sibley Memorial Hospital. Table 3-3 summarizes the expertise and the background of the expert panel, whose opinions were elicited in different stages of model building and model validation in this research. The panel includes experts from academia as well as private practice.

- years of clinical data from one of Harvard Medical School’s major teaching hospitals

Expert	Education	Expertise
Expert 1	MD, Surgeon	Faculty, Director, Clinical System Analysis (15+ years), Physician (23 years)
Expert 2	MD, Surgeon	VP Patient Safety and Quality (2 years), Physician (36 years)
Expert 3	MHA, RN	Director, Quality Improvement and Risk Management (6 years), Nurse (30 years)
Expert 4	MD, PhD	Faculty, Critical Care Medicine (14 years)
Expert 5	MD	Risk management, CMO (1 year), Physician (20 years)
Expert 6	MD	Pathologist, Nuclear Medicine, Risk Management (4 years)
Expert 7	MD, Surgeon	Patient Safety, Physician (20+ years)
Expert 8	MD	Faculty, Internal Medicine and Residency Program Director, Physician (25 years)
Expert 9	MD	Attending physician, Oncologist (8 years)
Expert 10	MD, PhD	Faculty, Director, Quality and Safety Research Group
Expert 11	MD	Deputy Director, National Clinical Public Health Program, Physician (4 years)
Expert 12	MBBS	Patient Safety
Expert 13	MM	Anesthesiologist with Expertise in Formal Risk Analysis
Expert 14	MD	Pediatric Anesthesiologist
Expert 15	LSW	Hospital Director of patient Safety and Risk management, Former Clinical Social Worker
Expert 16	MD	Hematologist
Expert 17	MD	Primary care

Table 3-3. Panel of experts

Some experts from this panel contributed only to parts of the modeling and quantification process, and some of the experts were involved in all stages of the modeling and quantification.

3.5.7 Eliciting the Structure of the Models; Qualitative Part

To construct the structure network (i.e.) for each of the Bayesian networks for specific hospital acquired adverse events that are studied for this research, namely Pressure Ulcer and Vascular catheter-Associated Infection (Line Infection) , and also the system dynamics part of the model, we started off by the factors existed in the literature and with the help of one the experts we drafted the sketch of the model. This first draft was then taken to each of the experts and was discussed with them in face-to-face interviews. The interview process was carried out in 3 different phases:

Phase One:

First the big picture of the research was presented. This included:

1. The methodology that was used and the combination of system dynamics and Bayesian belief network formalism.
2. Our hypothesis that healthcare organization's decisions in response to unfavorable revenue gap to reduce costs and close the gap may in the long run affect the risk of adverse events in the hospital and ultimately increase the costs in many ways. Each of the formalisms (SD and BBN) was explained through several examples.

Each of the individual sub modules; a system dynamic model to address organizational level decisions and two Bayesian networks to depict a causal model for two specific adverse events (PU and Line infection), were then discussed in several interviews.

Phase Two:

In the second phase the first draft of each individual model was discussed, that is 3 interviews for 3 sub modules (the system dynamics model, the pressure ulcer BBN and the line infection BBN) were conducted at this phase. This phase of interview included the following steps:

1. Giving a brief introduction/mind refresher of the problem under study and the tool. For instance, if pressure ulcer BBN was the subject of the interview , a brief introduction to Bayesian belief networks (e.g. how they are constructed, what type of problems they could solve, what the variables and probabilistic relations meant, the conditional probability tables etc.),was presented through several examples using the forms that can be found in appendix A&B.
2. Expert was asked to look the first draft model (built based on literature and one of the expert's opinion) and include, exclude, modify or edit in way, any variable or any causal relation between the variables. The analyst would record the expert's justifications on his/her modification to the first draft model.

This phase of the interviews is purely qualitative. The goal of this phase is to reach a model that experts agree on and believe is sufficiently representative of the problem under study.

Phase Three:

After collecting expert's opinions, their addition, deletion and modification to the first draft model, the analyst included all these modifications into the model, which resulted in an updated version of the model. This phase was done under the supervision of the expert who provided the original draft of the models.

Next, this updated version of the model was taken back to each individual expert and each expert was asked to review the structure of this version (with all experts' modifications and corrections included). In the case that experts still felt the need to make modifications, these modifications were discussed with other experts as well and after reaching a consensus was finalized into the model. This phase of expert interviews may have been done in more than one interview session. Then the experts were asked to rate the model on the scale of 0 to 100, in each of the following categories:

1. *Completeness.* From your perspective, to what extent does this model capture all important and relevant phenomena for the particular problem that we are studying? On a scale from 0 to 100, 0 would correspond to a model that does not include some important and relevant details, whereas 100 would correspond to a model that includes all details that you consider important. What number would you assign?

2. *Accuracy*: From your perspective, how accurately or realistically does the model depict important factors that influence risk of experiencing pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that is unrealistic, over-idealized or inaccurate, whereas 100 would correspond to a model that is realistic and accurate. What number would you assign?
3. *Ease of understanding*: From your perspective, how easy is it to understand the overall logic of the model. On a scale from 0 to 100, 0 would correspond to a model that is difficult to follow, even with extensive explanation, and a 100 would correspond to a model that is readily understandable. What number would you assign?
4. *Perceived predictive validity*: From your perspective, if you were to use this model, how well could you predict the risk of pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that does not help at all with predicting effects, and a 100 would correspond to a model that predicts the effects very well. What number would you assign?

The final product of these three phases, were models that were qualitatively verified and validated by experts and were ready for quantification. Phase four of the interviews was concerned with eliciting the parameters of the model, which will be discussed in the next section.

Figure 3-7, summarizes the phases of model construction and the validation of the qualitative part of the models.

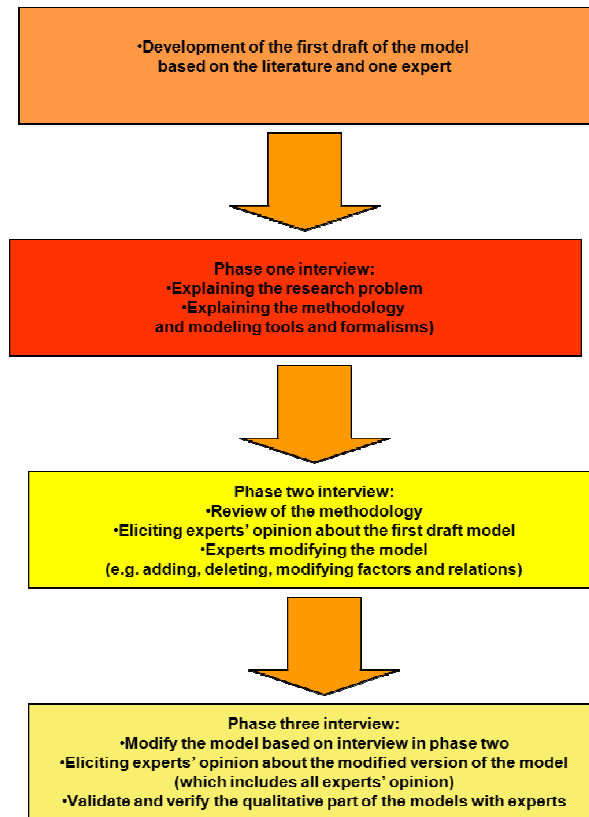


Figure 3-7. Model construction and qualitative validation process

It is worth mentioning that for elicitation of the structure of the models, for each sub model, each expert took part in four 30 minutes interview sessions. That is each expert on average spent about 6 hours in the course of 12 months, on the construction and validation of the qualitative (structure) part of the models (Pressure Ulcer BBN, Line Infection BBN and the SD model). This is excluding the expert providing the first draft

of the models, who dedicated on average half an hour of her time per week during this period.

3.5.8 Eliciting the Parameters of the Model (Model Quantification)

Expert judgment techniques are useful for model quantification where for various reasons including cost, uniqueness of the situation under study, difficulties and other reasons, none or not enough observations have been made in order to quantify data with real observed data.

3.5.9 Formats of Elicitation

Expert's opinion on the quantity of interest can be elicited in different ways and forms.

A. Direct Elicitation

In this form of elicitation, we elicit a direct estimate of expert's degree of belief on the issue under the study, which simply involves asking the expert to state his or her response and degree of belief on the subject. Different approaches and formats of elicitation may fall in this category. For instance, Response Scale; where experts choose between ranges of feasible responses presented to them, also is another method in this category (O'Hagan, 2006).

Although direct elicitation is the most straightforward method of elicitation, some concerns in the literature have been raised about the reliability of the results of this method. Especially when probabilities are being elicited and from experts who are not

quite familiar with the notion of probabilities (Ayyub, 2001). Educating the experts with the basic concepts of probability, and finding an efficient way for asking the questions from experts, in a way that is closer to expert's day to day experience may help alleviate this problem. For example if probability of the disease X is being elicited, instead of asking the experts "what is the probability that a patient will develop disease X?" it may be more efficient to ask "if you have 100 patients, how many of them would develop X, to your opinion". In other words, asking the questions in terms of relative frequency rather than probability.

B. Indirect Elicitation

The indirect method is based on betting rates, in order for the experts to reach to a point that they are indifferent between the options that are presented to them. For instance, if you are presented with an opportunity to win 100\$, and have an option to bet on event A or bet on throwing a "1" on a dice, which would you pick? If you pick betting on event A, it shows that your subjective probability of event A is greater than $\frac{1}{6}$. A sequence of bets may be used to refine and specify the subjective probabilities of experts more precisely (O'Hagan 2006, Ayyub, 2001).

3.5.10 Challenges and Generic Issues in Eliciting Expert Opinion

Besides the task of selecting a group of experts that are able and willing to contribute to the elicitation process, which could turn out to be quite a demanding task, other issues in elicitation may also be of concern. One such challenge is the issue of "biases".

Expert opinion is subject to biases; that is the possibility of overestimation, underestimation and overconfidence. Experts may provide their opinion with more certainty than is justified with their knowledge on quantities being assessed. Overconfidence is especially more common in assessing confidence intervals on an estimated quantity (Ayyub, 2001). Biases appear at many levels. Discussions in the literature (Bedford and Cooke, 2001, Otway and von Winterfeldt, 1992) could be found on mindset (unstated assumptions used by experts), structural biases (occurring through the level of detail in one parts of a study), motivational biases (when expert has a stake in the outcome of the study), cognitive biases (overconfidence for example), anchoring (when expert bases his or her opinion on an estimate given to him or her) and availability (when overestimates about events that can be recalled and underestimates about the events that are difficult to recall).

Even though these problems could not be entirely avoided, it is possible to guard against and control them, at least to some extent, by taking effective measures such as providing needed training to the experts and the use of calibration techniques (Ayyub, 2001).

Another challenge relating to quantitative expert elicitation, which could be seen in the literature is eliciting probabilities and the presentation format for communicating probabilities. For instance van der Gaag, et al. (2002), express that their experts had considerable difficulty understanding conditional probabilities using probability scales. For quantitative elicitation (including probabilities and conditional probabilities) we asked the experts for their opinion both in terms of probabilities and frequencies and fortunately no such difficulties were experience in communicating with the experts.

3.6 Validation and Verification

Model verification and validation (V&V) are essential phases of model development process. Verification is the process that ensures that the conceptual model has been translated into a computer model with no mistakes and with sufficient accuracy. Validation on the other hand, ensures that the model addresses the problem at hand with sufficient accuracy, and meets the intended requirements from the methodology and results perspective. In other words, with validating a model we want to ensure that the model addresses the problem of interest and provides sufficiently accurate information about the system being modeled. We are emphasizing the term “sufficiently accurate”, since no model of the real world is 100% accurate, but validation ensures sufficient accuracy with reference to the purpose the model is being used (e.g. demonstration models vs. others) (Robinson, 1997).

In Figure 3-8, Sargent (2004), shows how a verification and validation process needs to be involved in each step of the model building process, and also shows various forms of validations. There are many methods of verification and validation available to modelers, and unfortunately no study shows which are more effective and efficient, but below is a summary of some of the more common techniques (Robinson, 1997).

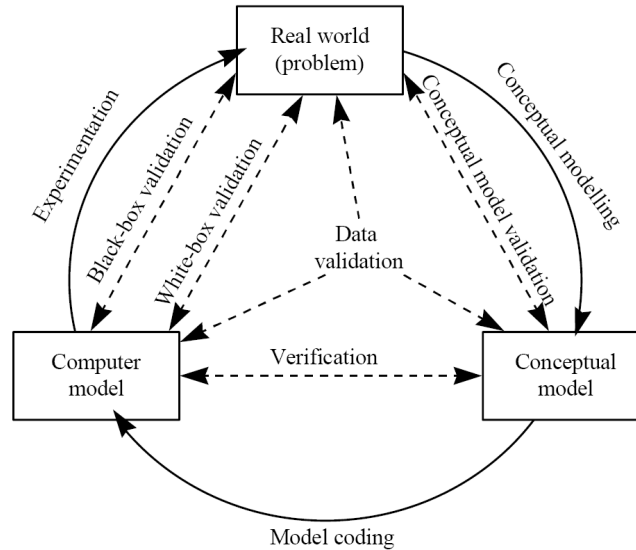


Figure 3-8. Simulation model verification and validation in the modeling process (Sargent, 2001)

- Conceptual Model Validation (i.e. is the level of detail in this model sufficient to answer the question at hand? are the assumptions correct and are all important variables included in the model),
- Data Validation (i.e. are data needed for model building and quantification accurate and reliable?),
- White-Box Validation (i.e. does each part of the model represent the real world with desired level of accuracy?),
- Black-Box Validation (i.e. does the overall model represent the real world accurately?)

Carson (2002) also proposes a simple rather intuitive framework for verification and validation;

- Testing the model for face validity , i.e. examining the model’s output measures of performance for a given scenario and determining how reasonable they are
- Testing the model over a range of input parameters, i.e. run sensitivity analysis and look for anomalies in the output
- Compare model predictions to past performance of the actual system

Much of what we have discussed in this section, is a reflection of simulation models’ validation literature, but in general could be applied to any type of model building activity. In validating and verifying the models in this work, we have applied this general framework. The methodology presented in this work includes a system dynamics formalism (a simulation model) and Bayesian belief network formalism (a probabilistic network). The V&V process used to validate the models in research is twofold; qualitative validation and quantitative validation. Since the models in both cases have been developed using subject matter experts’ input, much of the qualitative validation (both in system dynamics and Bayesian belief network models) are rather built in the model development process. The models have been developed and matured to the current version through much iteration in many rounds of interviews with as many as 11 experts.

The quantitative validation process in a nutshell, consists of using a few years of available data to build and calibrate the model and using data available for years other than the ones used for model building and calibration, to evaluate the performance of the models. More details are discussed on V&V for each of the models in chapter 4, where each model is discussed and the development and quantification steps are explained.

4 Model Development for Adverse Events; Pressure Ulcer

In this chapter we will discuss the BBN model developed for risk of hospital acquired pressure ulcer. The chapter provides a background on pressure ulcer, risk assessment tools used to assess the risk of pressure ulcer, and finally development, quantification and validation of the pressure ulcer BBN.

4.1 Introduction

A Pressure Ulcer (PU) is a skin break that does not heal and often causes irritation. Heels, elbows and buttocks areas of the body are most at risk. As the National Pressure Ulcer Advisory Panel (NPUAP) defines it “Pressure Ulcer (PU) is a localized injury to the skin and/or underlying tissue usually over a bony area, as a result of pressure in combination with shear and/or friction”. The NPUAP further categorizes the severity of PUs in the following stages in Table 4-1:

Stage 1	Intact skin with non-blanchable redness of a localized area usually over a bony prominence. Darkly pigmented skin may not have visible blanching; its color may differ from the surrounding area
Stage 2	Partial thickness loss of dermis presenting as a shallow open ulcer with a red pink wound bed, without slough. May also present as an intact or open/ruptured serum-filled blister
Stage 3	Full thickness tissue loss. Subcutaneous fat may be visible but bone, tendon or muscles are not exposed. Slough may be present but does not obscure the depth

	of tissue loss. May include undermining and tunneling
Stage 4	Full thickness tissue loss with exposed bone, tendon or muscle. Slough or eschar may be present on some parts of the wound bed. Often include undermining and tunneling

Table 4-1. Different stages of pressure ulcer

Though pressure ulcers are potentially preventable, unfortunately they present a common condition especially among high-risk population such as elderly and patients with impaired physical mobility (Reddy, Gill and Rechon, 2006).

In the United States, studies suggest that in acute care the prevalence of pressure ulcer ranges from 3.5 to 29% (estimated at 15% by NPUAP) (Ayello and Barden, 2002) , 2.2 to 26% among those in long-term care and 10 to 17% in homecare (Reddy, et al., 2006). Some studies suggest that the prevalence figures in spinal units are as high as 50% [Keller, et al. 2002]. Literature also suggests similar prevalence statistics in European hospitals (Papanikolaou, et al., 2007).

Pressure ulcers are painful for patients and costly to care for. An estimated 1.3 to 3 million pressure ulcers are treated in U.S. hospitals every year with an estimated cost of \$500 to \$40,000 to heal each ulcer [Lyder, 2003,] and may even cost up to \$ 75,000 per patient [Keller, et al. 2002, Reddy et al., 2006]. U.S. expenditures on treating pressure ulcers have been estimated to be \$11 billion each year. This number in the UK has been estimated in a 1993 study to be in the range £180-£231 million, which accounts for 0.4 -

0.8% of their health spending (Bennett, Dealey and Posnett, 2004). The development of pressure ulcers may also indicate neglect and mismanagement and have legal implications; 87% of litigation settlements regarding pressure ulcers in long-term care (LTC) settings have been in favor of LTC residents [Reddy, et al., 2006]. If pressure ulcers are to be prevented and the risk of PUs is to be controlled and reduced it is essential to identify patients who are at risk of experiencing this adverse event. Moreover, a range of preventive measures including use of pressure reducing mattresses and patient repositioning are available -even though limited information on their effectiveness exists- (Baldi, et al., 2010) but before any prevention plans are put in place, some form of risk assessment of individual patient's chances of PU should be carried out (Papanikolaou, et al., 2007, Borlawsky, 2004). Though some clinicians may believe that performing an informal PU risk assessment would suffice, research has shown that when a formal risk assessment is not undertaken, clinicians have consistently tended to intervene only at the highest levels of risk of PU, leaving many patients susceptible to the risk of hospital acquired pressure ulcer. It has also been shown that in studies where formal risk assessment was performed and preventive measures were taken accordingly, the incidence of PUs had dropped by 60%, with decreased severity of PUs and cost of care [Ayell and Braden, 2002]. In the next session, some of the more popular PU risk assessment tools are reviewed.

4.1.1 Risk Assessment Tools

Since a comprehensive and detailed risk assessment of every individual patient's vulnerability to pressure ulcer, based on the principals of wound healing, requires gathering a vast amount of knowledge and may become practically impossible. Several

risk assessment tools or risk assessment scales (RAS) have been designed since the 1960's, as a shortcut to produce a quick assessment and help practitioners identify patients who are at risk of developing pressure ulcer. Current guidelines underline that RASs should be used as an addition to provider's clinical judgment and not as a replacement. To date over 20 of such scales are described in the literature (Papanikolaou, et al., 2007). These tools include, among others, the Norton scale [Norton, McLaren, Exton-Smith, 1962], the Gosnell scale (Gosnell DJ., 1973). , the Braden scale (Bergstrom, Braden, Laguzza, Holman 1987), the Waterlow scale (Waterlow, 1985). Some of these scales such as Norton's and Waterlow's have been developed in Europe and others were created in the United States.

Typically, these scales produce assessments of a set of internal and external factors (e.g. mobility, nutrition, etc.) that are generally believed to be contributing factors in development of pressure ulcers. A numerical value is assigned to each of these factors based on patient's conditions, and these values are then summed to create a total score. The total score is usually compared to a critical value or a cutoff point, and hence it is used as an indication of patient's susceptibility to experiencing pressure ulcers.

Keller [2002] has summarized the risk factors considered by some of the well-known risk assessment scales.

Factors	Norton	Gosnell	Andersen	Waterlow	CBO	Douglas	Braden	Pressure Sore Prediction Score
Neurology					+			
Sensory perception							+	
Activity	+	+				+	+	+
Mobility	+	+	+	+	+		+	+
Moisture								+
Friction							+	
Nutrition		+					+ ^a	+
Physical condition	+					+		+
Mental state	+	+	+		+	+		+
Incontinence	+	+	+	+	+	+		+
Weight			+	+				
Skin state			+	+				
Gender				+				
Age			+	+	+			
Appetite				+				
Special risks				+ ^b	+ ^c	+ ^d		
Pain							+	
Dehydration			+					
Temperature					+			

Figure 4-1. Pressure ulcer risk factors (Keller, 2002)

Between the above mentioned risk assessment tools, the Braden scale is perhaps the most widely used in the United States. As a representative of this set of RASs, the Braden scale is discussed in the next session.

4.1.1.1 The Braden Scale

Following an observation that despite nursing staff's attention to repositioning and care of the skin of nursing home's patients in the US, poor nutritional condition was a major contributor to the formation of pressure ulcer, the Braden scale was developed in the 1980's to assess the susceptibility to the risk of pressure ulcer [Papanikolaou, et al., 2007, and Braden et al., 1987]. The Braden risk assessment tool is a linear combination of six risk indicators, formally shown as:

$$S_B = Z_1 + Z_2 + Z_3 + Z_4 + Z_5 + Z_6$$

Where the parameters are:

Z_B : The Braden score which represents the risk of developing pressure ulcer;

Z_1 : Sensory perception

Z_2 : Activity score

Z_3 : Mobility score

Z_4 : Nutrition score

Z_5 : Moisture score

Z_6 : Friction and shear score

Factors Z_1 to Z_3 , sensory perception, activity and mobility cover the clinical situations that expose patients to intense and prolonged pressure. Factors Z_4 to Z_6 , Nutrition, moisture and friction and shear cover the conditions that have an adverse effect on skin's tolerance for pressure. Given that the nurses have received proper training, they can provide necessary preventive interventions based on an individual patient's needs determined by the Braden score [Papanikolaou, et al., 2007, Ayello and Braden, 2001]. Figure 4-2, shows a formal worksheet for assessing a patient's risk for developing pressure ulcer using the Braden scale.

Each of these subscales is scored from 1-3 or 4, for total scores that range from 6-23. A lower Braden scale score indicates a lower level of functioning and, therefore, a higher level of risk for pressure ulcer development. A score of 19 or higher, for instance, would

indicate that the patient is at low risk, with no need for treatment at this time. This is based on the initial suggested critical score for the Braden scale, $S_B < 16$, at which skin breakdown, was thought to be commenced. This cut off point has since been disputed, for instance Bergquist and Frantz (2001) [Bergquist and Frantz, 2001, Papanikolaou, et al., 2007] have suggested 19 as the cut off score. It has also been suggested that it may be more efficient for healthcare units to determine their own critical point, considering the needs of their patient population and local clinical settings.

BRADEN SCALE FOR PREDICTING PRESSURE SORE RISK

Patient's Name _____	Evaluator's Name _____	Date of Assessment _____				
SENSORY PERCEPTION ability to respond meaningfully to pressure-related discomfort	1. Completely Limited Unresponsive (does not react, flinch, or grasp) to painful stimuli, due to diminished level of consciousness or sedation OR limited ability to feel pain over most of body.	2. Very Limited Responds only to painful stimuli. Cannot communicate discomfort except by moaning or restlessness OR has a sensory impairment which limits the ability to feel pain or discomfort over 1/2 of body.	3. Slightly Limited Responds to verbal commands, but cannot always communicate discomfort or the need to be turned OR has some sensory impairment which limits ability to feel pain or discomfort in 1 or 2 extremities	4. No Impairment Responds to verbal commands -has no sensory deficit which would limit ability to feel or voice pain or discomfort.		
MOISTURE degree to which skin is exposed to moisture	1. Constantly Moist Skin is kept moist almost constantly by perspiration, urine etc. Dampness is detected every time patient is moved or turned.	2. Very Moist Skin is often, but not always moist. Linen must be changed at least once a shift	3. Occasionally Moist Skin is occasionally moist, requiring an extra linen change approximately once a day	4. Rarely Moist Skin is usually dry, linen only requires changing at routine intervals		
ACTIVITY degree of physical activity	1. Bedfast Confined to bed.	2. Chairfast Ability to walk severely limited or non-existent. Cannot bear own weight and/or must be assisted into chair or wheelchair.	3. Walks Occasionally Walks occasionally during day, but for very short distances with or without assistance. Spends majority of each shift in bed or chair.	4. Walks Frequently Walks outside room at least twice a day and inside room at least once every two hours during waking hours.		
MOBILITY ability to change rank to the body position	1. Completely Immobile Does not make even slight changes in body or extremity position without assistance.	2. Very Limited Makes occasional slight changes in body or extremity position but unable to make frequent or significant changes independently.	3. Slightly Limited Makes frequent though slight changes in body or extremity position independently.	4. No Limitation Makes major and frequent changes in position without assistance.		
NUTRITION usual food intake pattern	1. Very Poor Never eats a complete meal. Rarely eats more than 1/2 of any food offered. Eats 2 servings or less of protein (meat or dairy products) per day. Takes fluids poorly. Does not take a liquid dietary supplement. OR is NPO and/or maintained on clear liquids or IV's for more than 5 days.	2. Probably Inadequate Rarely eats a complete meal and generally eats only about 1/2 of any food offered. Protein intake includes only 3 servings of meat or dairy products per day. Occasional will take a dietary supplement. OR receives less than optimum amount of liquid diet or tube feeding.	3. Adequate Eats over half of most meals. Eats a total of 4 servings of protein (meat, dairy products) per day. Occasionally will refuse a meal, but will usually take a supplement when offered. OR is on a tube feeding or TPN regimen which probably meets most of nutritional needs.	4. Excellent Eats most of every meal. Never refuses a meal. Usually eats a total of 4 or more servings of meat and dairy products. Occasionally eats between meals. Does not require supplementation.		
FRICTION & SHEAR	1. Problem Requires moderate to maximum assistance in moving. Complete lifting without sliding against sheets is impossible. Frequently slides down in bed or chair, requiring frequent repositioning with maximum assistance. Spasmodic, contractures or agitation leads to almost constant friction.	2. Potential Problem Moves freely or requires minimum assistance. During a move skin probably slides to some extent against sheets, chair, restraints or other devices. Maintains relatively good position in chair or bed most of the time but occasionally slides down.	3. No Apparent Problem Moves in bed and in chair independently and has sufficient muscle strength to lift up completely during move. Maintains good position in bed or chair.			
					Total Score _____	

© Copyright Barbara Braden et al Nancy Bergstrom, 1988 All rights reserved

Figure 4-2. Braden scale risk assessment worksheet

4.1.1.2 Validity and Reliability

Given the number of available RASs, the question may be raised, whether and why one scale may be preferred over another. Literature suggests the clinician should decide to use a scale by examining its reliability and validity [Ayello and Barden 2001]. What is meant by reliability here is consistency, i.e. the degree of agreement among raters (inter-rater-reliability). A common measure of reliability for a RAS is percentage agreement or the percentage of cases in which different clinicians/raters assign the same score to the same patients. Validity or accuracy on the other hand is the RAS's ability in correctly predicting whether a patient will develop pressure ulcer. Predictive validity is twofold; sensitivity and specificity. Sensitivity is the percentage of patients who do develop a pressure ulcer and were indeed identified as patients 'at risk' by the RAS. Therefore good sensitivity for a risk assessment tool means correctly identifying "true positives" with minimum "false negatives". Specificity is the percentage of patients who do not develop a pressure ulcer and were in fact identified as patients 'not at risk' by the RAS. Hence good specificity for a risk assessment tool means correctly identifying "true negatives" with minimum "false positives".

Even though the Agency for Healthcare Research and Quality (AHRQ) guidelines has mentioned two of these RASs (the Braden Scale and the Norton Scale) to be appropriate tools in assessing the risk of pressure ulcer due to the larger number of clinical research in support of their reliability and validity and having received the most clinical attention (Smith , 95) –although some studies have argued otherwise about their effectiveness (Defloor and Grypdonck, 2004)-, unfortunately the validity and reliability of many of the pressure ulcer risk assessment scales are questionable and no general agreement exists

with respect to the usefulness of these scales (Keller, 2006, Pandorbo-Hidalgo, et al., 2005). Many studies have reviewed these scales and a good number of them have examined the predictive validity of these risk assessment scales and have reported substantial variations in the predictive validity both within the same scale and across different scales when used in different health care settings and/or different patient populations (Papanikolaou, et al., 2007). For studies on the validation of some of these scales please see (among many): Spera et al, 2010, Seongsook, et al. 2003, Defloor and Grypdonck, 2004...

4.1.1.3 Pitfalls of Scoring Approach to Risk Assessment

Despite the fact that using risk assessment tools, in addition to clinician's judgment, provides some useful information in identifying the patients at risk in developing pressure ulcer and helps practitioners make an informed decision in implementing appropriate preventive interventions, there are methodical shortcomings that are common between these RASs.

In the scoring system that is used in these risk assessment scales to identify patients at risk and patients not at risk, every risk factor contributes equally to the overall risk score. In other words the scoring approach to risk assessment assumes that all the factors have equal effect on the overall risk of developing pressure ulcer. The equal-weighting approach while being the simplest way to scale scoring, fails to recognize that some factors may play a more significant role and therefore should have a larger contribution to

the overall risk score (Papanikolaou, et al., 2007). For a more accurate predictive measure, the magnitude of the effect of each of the risk factors on the overall risk of developing pressure ulcer has to be considered, based on the importance that these factors empirically demonstrate. Failing to do so may project unrealistic risk scores that could possibly influence the effectiveness of the interventions and affect the allocation of resources.

Another rather important deficiency of most of risk assessment scales is that the effect of all risk factors contributes linearly to the overall risk score. This completely overlooks the fact that a certain factor in presence of other factors may, for instance, exponentially increase the risk of pressure ulcer. For example, consider the Braden scale. Given that Sensory and Nutrition are influencing factors in risk of pressure ulcer but the magnitudes of this influence is in a) presence of impaired mobility and b) un-impaired mobility could be very different.

4.2 Pressure Ulcer BBN Development

To assess the risk of developing pressure ulcer as a function of individual patient's risk factors and patient-provider (i.e. intervention related), a Bayesian Belief Network framework has been chosen. Use of BBNs in modeling the risk of experiencing pressure ulcers, not only alleviates the major criticism to the scaling risk assessment approach, namely the equal weighting of the risk factors, but also offers capabilities that could possibly provide more realistic, relevant and meaningful assessments;

- Since we construct the Bayesian Network based on the conditional probabilities, no equal weighting of the factors is assumed. Based on the

importance of each factor and the strength with which these factors influence the risk of pressure ulcer (obtained from field data and also expert judgment) we can determine the conditional probability that a patient will experience pressure ulcer given the states of all the risk factors.

- Using BBNs enables the analyst, to take into the account the fact that the degree of influence of one factor in risk of pressure ulcer may be different given the presence or absence of other risk factors.
- Bayesian Belief Networks are probabilistic in nature and the uncertainty of our assessment of pressure ulcer risk, given the state of all relevant risk factors can be expressed explicitly.

A Bayesian Belief Network, that includes or reflects the factors introduced in literature as factors influencing risk of pressure ulcer, has been developed. Additionally factors that the panel of experts thought to be of importance, and missing from the current risk assessment scales, have also been included. Figure 4-3 depicts this BBN. The validation process of the model has been detailed in section 4.4.

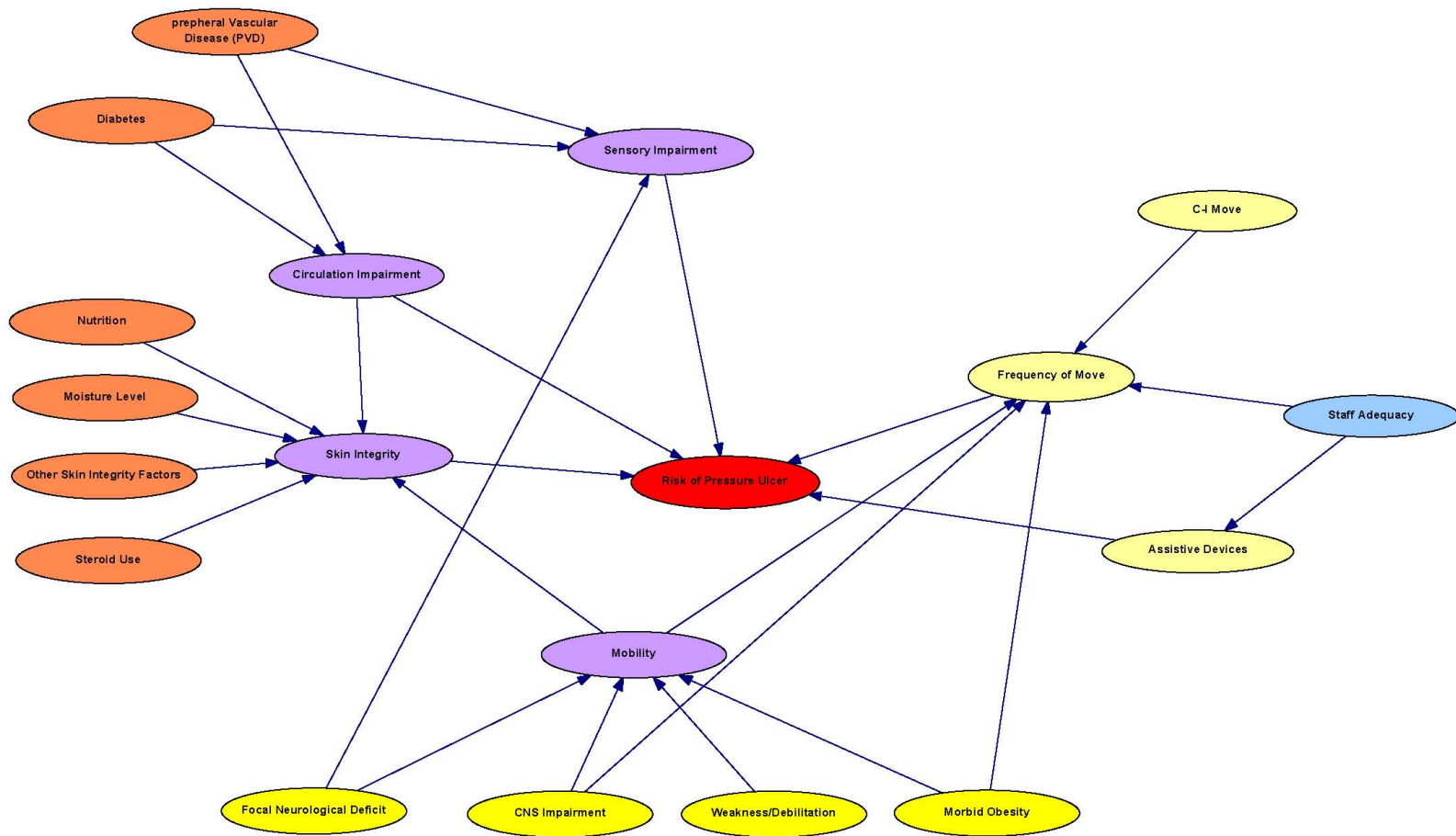


Figure 4-3. Pressure ulcer BBN

a. Circulation Impairment:

Poor blood circulation makes patients more susceptible to pressure ulcer. Although impaired circulation can be resulted from various conditions, in this model we have considered two major factors that may result in impaired circulation; diabetes and peripheral vascular disease (PVD).

This is a binary factor in the BBN and the possible states are Impaired and Unimpaired.

b. Peripheral Vascular Disease (PVD):

Peripheral vascular disease refers to diseases of blood vessels located outside the heart and brain. It is a circulatory problem in which narrowed vessels reduce blood flow to the legs, arms and kidneys [American heart association, www.americanheart.org].

This is a binary factor in the BBN and the possible states are PVD present and PVD absent.

c. Sensory Impairment:

Sensory impairment refers to a defect in sensing or passing on the impulse, which affects patients' ability to respond to pressure related pain and discomfort. Factors that may affect sensory impairment include diabetes, peripheral vascular diseases and focal neurological deficit.

This is a binary factor in the BBN and the possible states are Impaired and Unimpaired.

d. Skin Integrity:

Skin integrity is a description of whether or not patient's skin is intact. A number of conditions/factors may affect the integrity of skin, which include: Nutrition (food intake), Moisture level (the degree to which skin is exposed to moisture), Steroid use, Mobility, and Circulation impairment.

This is a binary factor in the BBN and the possible states are Normal and Abnormal.

e. Mobility:

Mobility refers to patient's ability to change and control his/her body position. In the BBN this is a binary node with states Impaired and Unimpaired mobility. Mobility is generally considered the most important risk factor in developing pressure ulcer and a necessary condition [Allman, et al., 1995, Lindgren, et al. 2004]. Factors affecting an individual's mobility impairment include: focal neurological deficit, central nervous system impairment, weakness and debilitation and morbid obesity.

f. Frequency of Move:

Another important factor in risk of developing pressure ulcer is whether the patient is being moved to different body positions frequently enough, especially when patient's own ability to move and mobility is impaired. This node reflects whether the staff can/do move the patient as often as the patient should be repositioned -it is important to note that detecting and preventing pressure ulcers systematically is labor intensive (Perneger, et al., 1998)- in order to reduce the risk of developing ulcer. This is also a binary node with

states: Adequate frequency of move and inadequate frequency of move. Adequacy of frequency of move on the other hand is influenced by staff adequacy (whether or not we have enough staff at the time to be able to frequently reposition the patient), C-I Move (Counter indication to move) (are there any limitations that may prevent the staff from moving the patient, for instance a patient recovering from open heart surgery) and morbid obesity (the heavy weight of the patient may make it extremely difficult for the staff to move the patient).

Currently, there is no empirical evidence to show the optimal frequency of repositioning the patient and it should be done based on patient's need, also taking into the account the surface upon which the patient is lying or sitting (Gunningberg, 2005).

g. Assistive Devices:

To relieve pressure, several strategies may be used including manual repositioning of the patients, which is discussed in "Frequency of Move" node of the BBN, and also use of assistive devices. These assistive devices include support surfaces such as cushions, mattress overlays, replacement mattresses or pressure relieving beds (Nixon et al., 2006), which reduce the risk of pressure ulcer (Reddy, et al., 2006, McInnes, et al., 2010). EUAP suggests that a patient receives appropriate preventive measures while in a chair or a bed if he or she is allocated one of the following (Gunningberg, 2005):

1. A powered device (i.e. with an electrical supply)
2. A non-powered device (i.e. low pressure foam mattress) and being repositioned every 2, 3 or 4 hours
3. No special device but being repositioned every 2 hours

The use of these assistive devices depends on their availability and also staff adequacy (whether or not the high level of workload prevents staff from providing patients with these devices).

4.3 Pressure Ulcer BBN Quantification

As discussed previously, building a Bayesian network for a certain application has three steps and involves three tasks. First, important variables and their possible states have to be identified. Second, the relationships between these variables are identified and are represented graphically with edges between the variables. The third phase is to obtain the numerical parameters, i.e. probabilities required for the quantification of the network from data or through domain expert elicitation (Druzdel and van der Gaag, 2000). The first two tasks that are concerned with establishing the structure of the network typically involve iterative and interactive sessions with domain experts. Multiple iterative cycles are required to revise the model(s), identify new variables and links or perhaps delete other variables and links and converge on a valid representation of the phenomenon that is being studied. For the first and the second task we have followed the process proposed by Marcot et al. (2006), for the peer review of the BBNs, we started with a basic influence diagram as base model and followed the peer review process using the panel of domain experts to develop, refine and validate the BBNs (both pressure ulcer and line infection BBNs). The process of peer review of BBNs has been discussed in detail in section 3.5.7. While the first and second tasks require moderate effort and time,

experience indicates that the third task, which is the elicitation of the quantitative information including the conditional probability table or CPT, requires the most effort (Kjaerulff and Madsen, 2008). Certain Modeling techniques are available to make the third task more manageable, without (or with minimum) compromising the accuracy of the results. We have used some of these techniques in quantifying the BBNs in this study such as parent divorcing and Noisy-OR gates which are explained in the next section.

4.3.1 Modeling Techniques

There are a number of modeling techniques and methods that could be used to simplify the specification of a Bayesian network. One of the reasons that these methods may be applied is to simplify the knowledge elicitation process. Kjaerulff and Anderson (2008), cover these methods and techniques in two categories:

- 1- Structure related techniques, that are used to adjust the structure of a probabilistic network
- 2- Probability distribution related techniques for the specification of conditional probability distributions, including techniques for capturing uncertain information and for reducing the number of parameters to be specified

Parent divorcing is technique from the first category, that reduces the number of probabilities to be assessed with making changes to the graphical structure of the BBN and Noisy-OR gates (and its generalizations) are an approach that falls into the second category which uses parametric probability distributions. Both of these techniques have been used in quantifying the BBNs in this research. Sections 4.3.1.1 and 4.3.1.2 explain these methods.

4.3.1.1 Parent Divorcing

Parent divorcing is a modeling technique that is commonly used to reduce the complexity of specifying and representing the effect of a large number of cause variables (parent nodes) on a single effect variable (child node) in a Bayesian network, with adjusting the structure of the network. The idea is to introduce an intermediate variable or a dummy node between the cause node (i.e. parent node) and the effect node (i.e. child node), such that the dummy node captures the impact of its parents on the child variable, in order to limit the size of the parent sets.

Figure 4-4, below depicts this idea through a simple Bayesian net. Child node “Y” has three direct parents X_1, X_2, X_3 . Applying the parent divorcing technique to Y and its direct causes results in creating a dummy variable “I” between Y and a subset of its parents X_1, X_2 , hence variable I will have X_1, X_2 as parents and variable Y will have X_3 and I as its parents.

If we assume binary states for all of the variables in this example, and also assume the conditional probability distribution for the original network as depicted in Figure 4-4

With the creation of dummy variable I, the BBN will change from figure 4-4.a to 4-4.c our conditional probability table will change to from figure 4-4.b to 4-4.d In other words instead of dealing with one distribution table of size 16 we create two tables of size 8 (Kjaerulff and Madsen, 2008). Figure 4-4.d top, shows the probability distribution for $P(I|X_1, X_2)$, and figure 4-4.d, bottom for $P(Y|X_3, I)$.

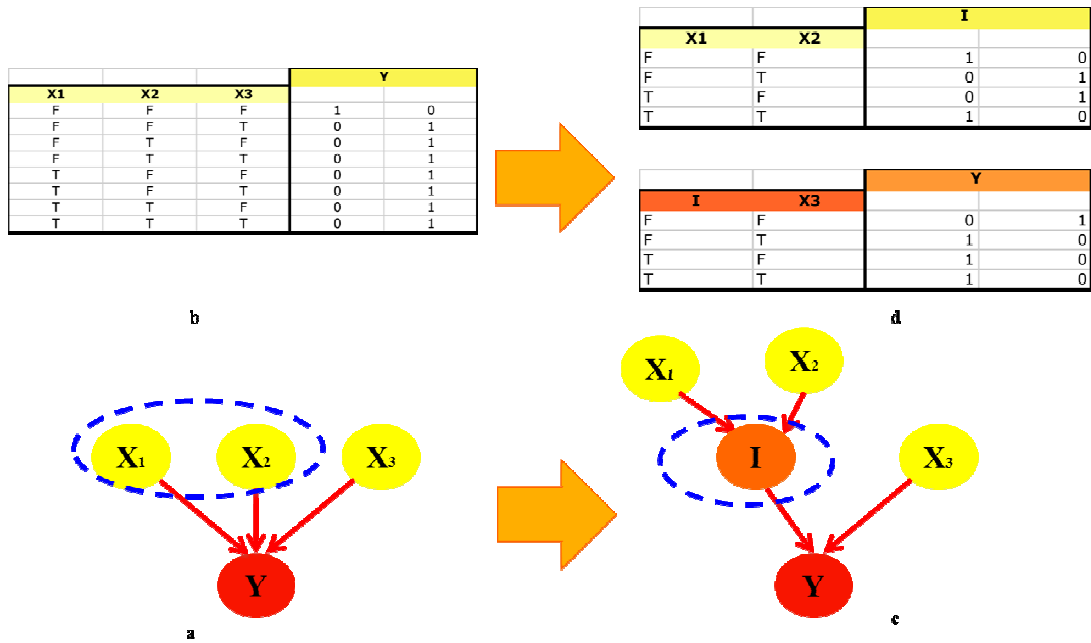


Figure 4-4. Parent divorcing

4.3.1.2 Noisy-OR Gates

In quantifying BBNs, some types of conditional probability distribution can be approximated with methods that require fewer parameters, and very often they approximate the true distribution sufficiently well while reducing the model building effort significantly (Onisko, et al., 2001). Noisy-OR gates by Pearl (1988) and their generalizations is one of such approaches. Noisy-OR gates are used to describe the interaction between causes X_1, X_2, \dots, X_n and their common cause Y . Two crucial assumptions are:

- X_i s are each sufficient to cause Y in absence of other causes
- X_i are independent of each other in causing Y

If each of the causes X_i , has a probability p_i of being sufficient to cause Y, then the Noisy-OR gate methods enables us to populate the entire conditional probability table (CPT), with only n parameters, p_1, p_2, \dots, p_n , where p_i is the probability that effect Y will be true if cause X_i is present and all other causes $X_j, j \neq i$ are absent. In mathematical representation;

$$p_i = pr(y | \bar{x}_1, \bar{x}_2, \dots, x_i, \dots, \bar{x}_n)$$

And the probability of Y given any subset X_p of causes X_i s, that are present will be:

$$pr(y | X_p) = 1 - \prod_{i: X_i \in X_p} (1 - p_i)$$

This formula is sufficient to derive the whole conditional probability of Y, conditioned on causes X_1, X_2, \dots, X_n (Pearl, 1988, Onisko, et al., 2001).

Extensions have been developed for the basic Noisy-OR gate such as Lemmer and Gossink (2004) propose a recursive Noisy-OR gate where the independence assumption of causes could be relaxed, and Henrion (1989) proposes a Leaky Noisy-OR gate for situations where the effect Y is true and all the causes X_1, X_2, \dots, X_n are absent. This extension could be used where a model is not capturing all the possible causes. In Leaky Noisy-OR gate, a parameter p_0 called the leak probability is introduced and its value is the combined effect of all causes of Y that are not modeled;

$$p_0 = pr(y | \bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$$

Which represents the probability that all causes, X_1, X_2, \dots, X_n , are absent but the effect Y is true. Henrion (1989) then derives the probability distribution of Y given a subset X_p of the X_i s which are present and the leak probability of p_0 as;

$$pr(y|X_p) = 1 - (1 - p_0) \prod_{i: x_i \in X_p} \frac{1 - p_i}{1 - p_0}$$

Diez (1993), also proposes an extension to Noisy-OR gates that includes multiple states for variables rather than binary states in the original Noisy-OR.

4.3.2 The Quantification Process

To proceed with the quantification of the Bayesian net for pressure ulcer the following steps in modification of the net, without compromising the causal structure of the Bayesian net and the accuracy of the output, have been taken.

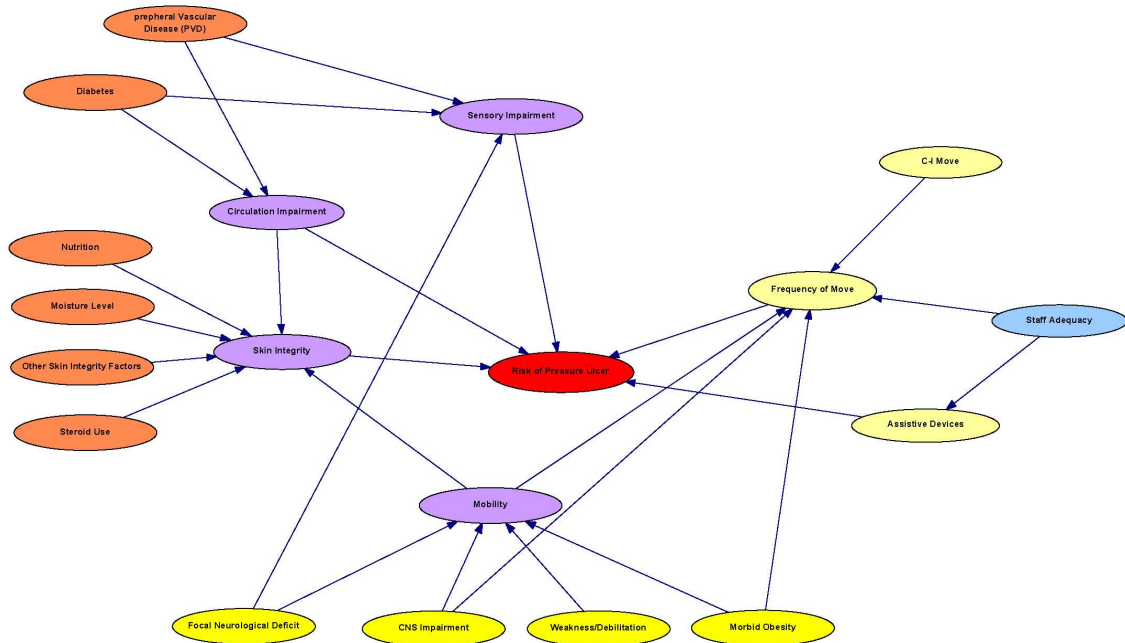


Figure 4-3, shows the pressure ulcer risk BBN, as it was qualitatively validated (factors and casual effects) by our panel of domain experts.

Data to establish the conditional probabilities was obtained by querying a clinical data archive at a large urban US medical center. This data repository contains diagnostic codes and clinical outcomes for 70,090 inpatients hospitalized over a 2-year period. After obtaining IRB approval, structured queries were constructed to identify conditions that were present in two distinct cohorts of patients: 1) patients who did not acquire a pressure ulcer during hospitalization and 2) patients who did acquire a pressure ulcer during hospitalization. At the time of discharge, expert codification of up to 15 physiological or disease condition codes are assigned to characterize the patient and the episode of care. Pharmacy and laboratory data for some of the patients analyzed in some cases either to confirm one or more diagnostic codes, disambiguate clinical conditions or identify additional patients in the cohort. As an example, in the case of the node “Skin Integrity”,

which may be affected by Nutrition, Moisture Level, Steroid Use, Circulation Impairment and other factors not specified in the model, we extracted all cases of compromised skin integrity due to any reason (specified or unspecified in the BBN model), to ensure that any patient with skin integrity issues is accounted for in this model. Similarly, for the nodes “Circulation Impairment” and “Sensory Impairment” the two most prevalent causing factors for these conditions as experts have identified are “Diabetes” and “Peripheral Vascular Diseases or PVDs”, among others (which data may or may not exist for). In the quantification of this model, we have identified all the cases of circulation impairment and sensory impairment among the hospital population regardless of the cause, to make sure that all patients with these conditions are included in the model. One important step was to distinguish those patients who acquired a pressure ulcer during their hospitalization from those who were treated for the condition, but had the condition at the time of admission to the hospital. To do this, we constrained the queries using a special ‘Present on Admission’ code that is used to classify patient conditions at this medical center.

Additionally, as it was explained in section 4.2, Mobility is a binary factor with states; Impaired Mobility =1, Un-impaired Mobility=0. The factors/conditions affecting a patient’s mobility are “Focal Neurological Deficit”, “CNS Impairment”, “Weakness /Debilitation” and “Morbid Obesity”. With the approval of experts, in quantifying this BBN we have assumed that if one of these causes is present then the patient will be considered to have mobility impairment. Since the factor “Mobility” is not readily available in datasets, to calculate the relative frequency of each state of the node “Mobility”, one can instead count the number of cases where at least one of the causes of

impaired mobility is present. This frequency divided by the total number of cases under study produces the relative frequency of states: “Impaired Mobility” and “Unimpaired Mobility”.

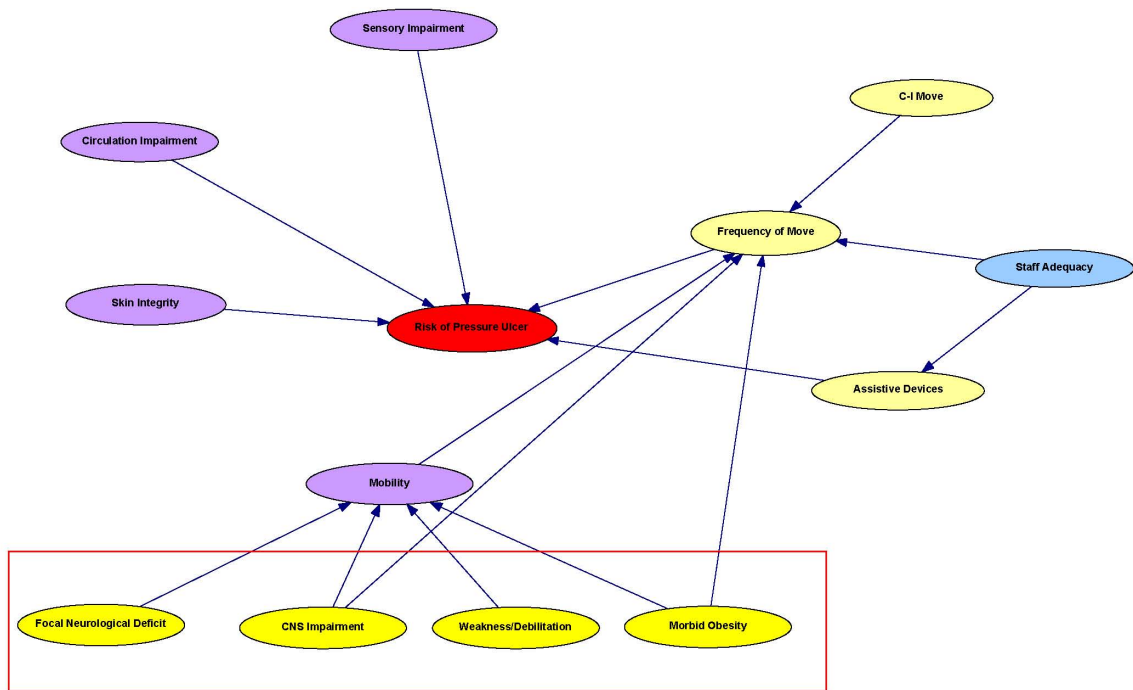


Figure 4-5. Pressure ulcer BBN, transformation 1

Hence, from data and quantification point of view the BBN will be transformed to Figure 4-5.

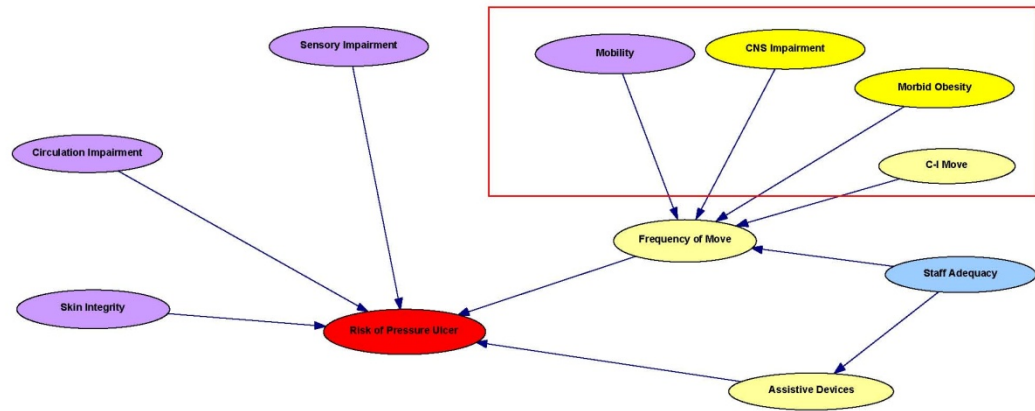


Figure 4-6. Pressure ulcer BBN, transformation 2

Frequency of Move is one of the most important and determining factors in a patient's risk of experiencing pressure ulcer. Unfortunately due to the difficulty in collecting information for this node and determining whether the patient's movement was adequate or not, this data does not exist in data bases [NOTE: some data exists, but it is unreliable, and there are a lot of false negative or β errors]. But we do have crisp data on the factors that experts think affect the frequency of move in a patient (e.g. mobility, CNS impairment, C-I move and Obesity (figure 4.6)). To deal with this situation and obtain the conditional probabilities of adequate and inadequate frequencies of movement and to capture the effect of four parent variables (Mobility, CNS impairment, Morbid Obesity and C-I Move) on the frequency of move, we are using the "parent divorcing technique" introduced in section 4.3.1.1, and creating a dummy node called "Aggregate Effect on Frequency of Move", a binary variable with the following states:

- 1) High: when at least one of the four factors (Mobility, CNS Impairment, Morbid Obesity or CI move) is present
- 2) Low: otherwise

The probability of Frequency of move being adequate or inadequate is then conditioned on this “Aggregate Effect on Frequency of Move” node and "Staff Adequacy". Figure 4-7 depicts this modification to the structure of BBN.

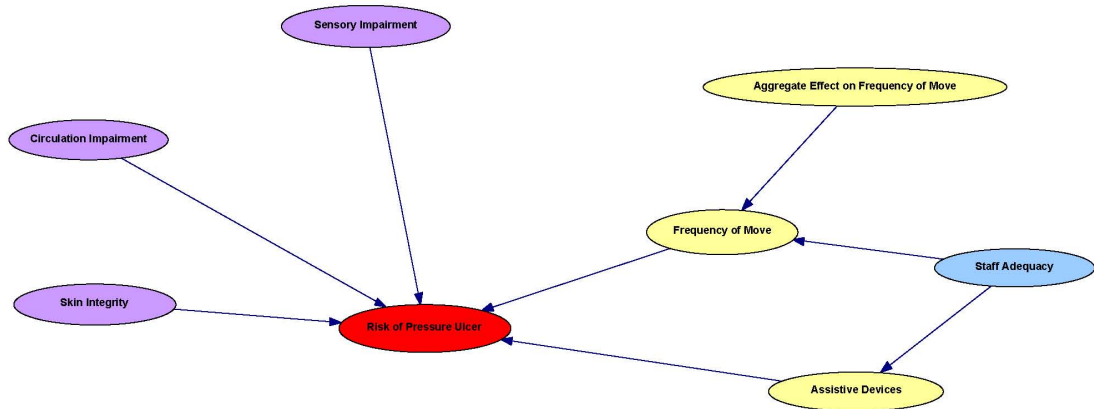


Figure 4-7. Pressure ulcer BBN, final transformation

4.3.3 The Conditional Probability Table (CPT)

Between years 2008 and 2010, we have used 70,090 patient records to construct the conditional probability table. Out of these patients a total of 149 patients had developed pressure ulcers while in the hospital. Since only hospital acquired pressure ulcers were of

interest in this study, we have started with the year 2008 patient data because only after this year whether the pressure ulcer was acquired while in hospital or not, was actually specified in data bases.

4.3.3.1 Marginal and Conditional Probabilities

After importing the data to Microsoft Access, the number of patients with condition(s) specified as risk factors in pressure ulcer BBN were counted, and relative frequencies of these factors and the conditional probabilities required to populate the conditional probability table and quantify the BBN were calculated as follows.

4.3.3.2 Marginal Probabilities

1. Skin Integrity

States:

- i. Skin Integrity Compromised =1
- ii. Skin Integrity Uncompromised=0

Total number of patients with Compromised skin integrity = 22905

Total hospital admissions = 70090

$$p(\text{Skin Integrity Compromised}) = \frac{22905}{70090} = 0.327$$

$$p(\text{Skin Integrity Uncompromised}) = 0.673$$

2. Circulation Impairment

States:

- i. Circulation Impaired = 1
- ii. Circulation Normal = 0

Total number of patients with impaired circulation = 29202

Total hospital admissions = 70090

$$p(\text{impaired Circulation}) = \frac{29202}{70090} = 0.417$$

$$p(\text{Normal Circulation}) = 0.583$$

3. Sensory Impairment

States:

- i. Sensory Impaired = 1
- ii. Sensory Unimpaired = 0

Total number of patients with impaired sensory = 17743

Total hospital admissions = 70090

$$p(\text{Sensory impaired}) = \frac{17743}{70090} = 0.253$$

$$p(\text{Sensory Normal}) = 0.747$$

4. Mobility

States:

- i. Mobility Impaired = 1
- ii. Mobility Unimpaired = 0

Any patient with Focal Neurological Deficit, Central Nervous System Impairment, Weakness/Debilitation or Morbid Obesity has been counted as a case of impaired mobility.

Total number of patients with impaired mobility = 26289

Total hospital admissions = 70090

$$p(\text{impaired Mobility}) = \frac{26289}{70090} = 0.375$$

$$p(\text{Unimpaired Mobility}) = 0.625$$

5. Aggregate Effect on Frequency of Move

States:

- i. High = 1
- ii. Low = 0

Any patient with Central Nervous System Impairment, Morbid Obesity, Impaired Mobility or Counter Indication to Move (C-I Move) has been counted as patient with high aggregate effect on Frequency of Move.

Total number of patients with high aggregate effect on frequency of move= 26608

Total hospital admissions = 70090

$$p(\text{High Aggregate Effect}) = \frac{26608}{70090} = 0.389$$

$$p(\text{Low Aggregate Effect}) = 0.621$$

6. Staff Adequacy

States:

- i. Adequate = 1
- ii. Inadequate=0

Data on staff adequacy is not properly reported and recorded. The available data for 627,595 patient cases used to quantify pressure ulcer BBN, indicates that in less than 6% of the cases staff adequacy was reported as being adequate and in over 94% of these cases data on staff adequacy was not recorded at all. Hence a more reliable estimate on the probability of “Staff Adequacy” would be obtained from subject matter experts. Some experts believed that this estimate would be different from one hospital to the other and the difference could be significant and pointed out that their estimates reflect their experience in their own institution. Table 4-2, shows experts estimates on the probability of staff adequacy. We will use these estimates as a prior, and update the probability of staff adequacy, with the staff adequacy probability obtained from the system dynamics part of the model, which calculates this probability as a function of patient complexity scores and the pressure to reduce operational costs, discussed in section 6.2.3.

Expert	Staff Adequacy	
	Adequate	In-adequate
Expert 1	0.95	0.05
Expert 2	> 90%	< 10%
Expert 3	80%	20%
Expert 4	85-90%	10-15%
Expert 5	95-98 %	2-5%
Expert 6	90%	10%

Table 4-2. Expert opinion; staff adequacy

4.3.3.2.1 Aggregating Experts Input

To aggregate experts' inputs which are given in the form of a point estimate, we will use a Bayesian framework, for treating non-homogenous data (Droguett and Mosleh, 2008, Droguett, 1999). The objective is to find $\phi(x)$ the population variability distribution of x (e.g. Probability of Staff In-Adequacy). To simplify matters we assume a parametric distribution for $\phi(x)$. Let $\bar{\theta} = \{\theta_1, \dots, \theta_m\}$ be the set of m parameters of $\phi(x)$, so that $\phi(x) = \phi(x|\bar{\theta})$. For instance, in the case of Lognormal distribution $\bar{\theta} = \{\mu_i, \sigma_i\}$ and:

$$\phi(x) = \phi(x|\mu_i, \sigma_i) = \frac{1}{x\sigma_i\sqrt{2\pi}} e^{-\frac{(\ln x - \mu_i)^2}{2\sigma_i^2}}$$

Uncertainty distribution over the space of $\phi(x)$'s, is the same as the uncertainty distribution over the values of $\bar{\theta}$. Also, for each value of $\bar{\theta}$, there exists a unique $\phi(x|\bar{\theta})$ and vice versa. Now our goal of estimating $\phi(x)$ reduces to estimating $\bar{\theta}$. Given the information/evidence available to us (denoted as E), in our case the estimates provided by our experts, and a prior distribution for $\bar{\theta}$, we can obtain an updated state of knowledge probability distribution over $\bar{\theta}$. That is,

$$\pi(\bar{\theta}|E) = \frac{L(E|\bar{\theta})\pi_0(\bar{\theta})}{\int_{\theta_1} \int_{\theta_2} \dots \int_{\theta_m} L(E|\bar{\theta})\pi_0(\bar{\theta})d\bar{\theta}}$$

Where;

$\pi_0(\bar{\theta})$ =Prior distribution of $\bar{\theta}$

$\pi(\bar{\theta}|E)$ =Posterior distribution of $\bar{\theta}$, given the information/evidence E (m-dimensional joint probability distribution over values of $\theta_1, \dots, \theta_m$)

$L(E|\bar{\theta})$ =Likelihood of information/evidence E given $\bar{\theta}$

The average distribution, of the distributions of x, is given by:

$$\bar{\phi}(x) = \int_{\theta_1} \dots \int_{\theta_m} \phi(x|\theta_1, \dots, \theta_m)\pi(\theta_1, \dots, \theta_m|E)d\theta_1 \dots d\theta_m$$

The expected value of $\bar{\theta}$ is given by:

$$E(\bar{\theta}) = \int_{\theta_1} \dots \int_{\theta_m} \bar{\theta}\pi(\bar{\theta}|E)d\bar{\theta}$$

Using the expected (average) value of $\bar{\theta}$ as the set of parameters of $\phi(x)$, we will obtain another point estimate of $\phi(x)$. In other words, $\phi(x|E(\bar{\theta}))$, is the distribution with the *mean value parameters*.

The likelihood function $L(E|\bar{\theta})$, is the probability of observing/eliciting the information

E, given that the set of the parameters of the population variability distribution is $\bar{\theta}$.

Assuming that data from individual sub population (i.e. estimates from each expert) are independent, the likelihood function can be written as the product of each sub population likelihood function:

$$L(X_1^*, X_2^*, \dots, X_N^* | X) = \prod_{i=1}^N L_i(X_i^* | X)$$

Where X_i^* is the i 'th expert's estimate.

To aggregate experts' opinion, using the Bayesian framework for non-homogenous data, discussed above, we are using version 1.5 of the R-DAT software (Prediction-Technologies.com).

The specification of the likelihood function depends on the type of information that is available. Expert-based likelihood that corresponds to the estimates of possible values of a quantity of interest (e.g. Probability of Inadequate Staffing) could be expressed with a lognormal likelihood model, and is specified in terms of median values and the analyst's confidence in terms of standard deviation or error factor values. In this case we are assuming an error factor of 2 for all of our experts.

Figure 4-8, shows the joint distribution of the parameters, of the distribution of "Probability of Inadequate Staffing", and Figure 4-9 shows the average distribution of the population variability distribution set, with mean 0.105 and variance 3.24 E-3.

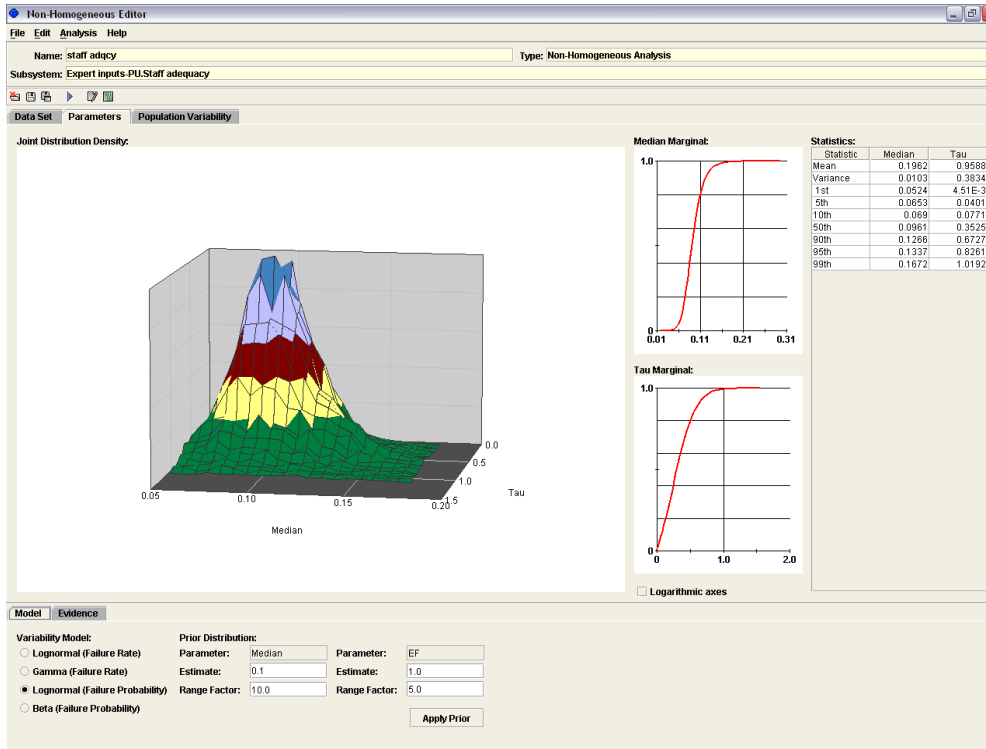


Figure 4-8. Joint distribution of the parameters, of the distribution of “Probability of inadequate staffing”

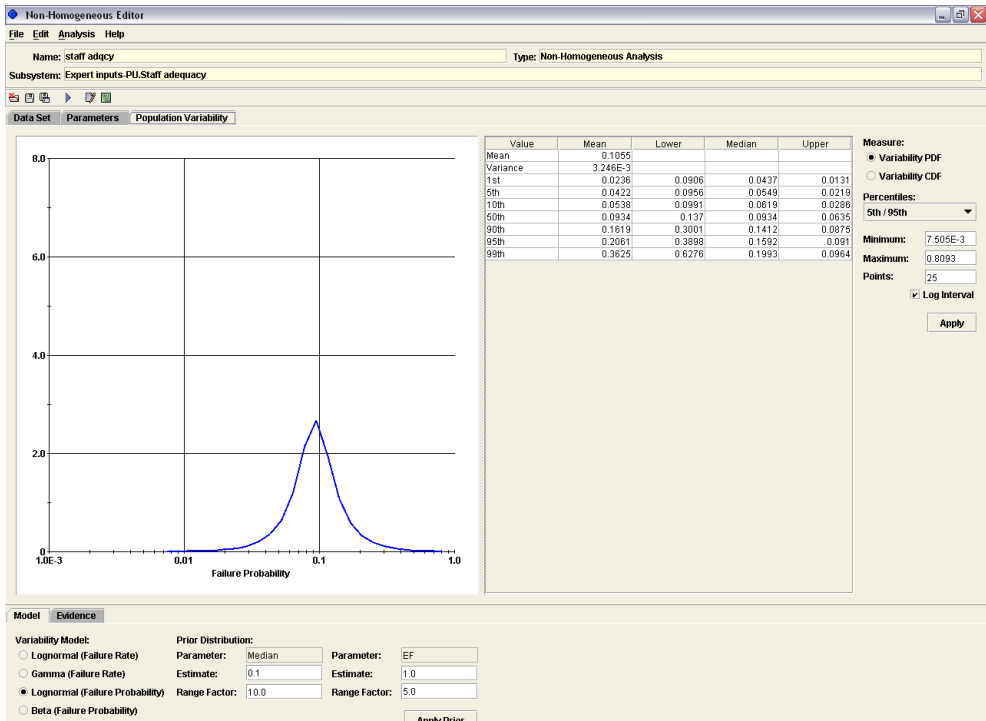


Figure 4-9. Average distribution of the population variability distribution set

4.3.3.2.2 Conditional Probabilities

7. Assistive devices

States:

- i. Used = 1
- ii. Unused = 0

The probability of using assistive devices such as pillows, cotton blankets etc., which expand weight-bearing surface (e.g. using pillows under the calf to elevate patient's heels off the bed surface) or reduce friction and/or shear is influenced by and conditioned on adequacy of staff. In other words do we have enough staff and does the workload allow them to utilize these devices that may reduce the risk of pressure ulcer. It should be mentioned that in the quantification of pressure ulcer BBN, we have assumed (with the approval of experts) that items such as pressure reducing mattresses and pressure reducing mattress overlays would be used automatically if they are available since they will ultimately be charged to the patient and hence would not be a variable in this BBN.

Unfortunately, data on usage of assistive devices is also not recorded as rigorously either and similar to the case of the node "Staff Adequacy" here in approximately 95% of cases no data at all was recorded and in 5% of cases data indicated that assistive devices were in use, and not a single record was found where no usage of assistive devices was reported. Once again experts' opinion would be a more reliable source in obtaining the probabilities (conditional) of usage of assistive devices given different states of the

variable “Staff adequacy”. The information experts were to provide answers to the following:

- 1- What is the probability of assistive devices being used, given we DO have adequate staff? i.e. : $p(\text{Assistive Devices} = 1 | \text{Staff Adequacy} = 1)$
- 2- What is the probability of assistive devices being used, given we do NOT have adequate staff? i.e.: $p(\text{Assistive Devices} = 1 | \text{Staff Adequacy} = 0)$
- 3- What is the probability of assistive devices NOT being used, given we DO have adequate staff? i.e.: $p(\text{Assistive Devices} = 0 | \text{Staff Adequacy} = 1)$
- 4- What is the probability of assistive devices NOT being used, given we do NOT have adequate staff? i.e.: $p(\text{Assistive Devices} = 0 | \text{Staff Adequacy} = 0)$

To ask experts these questions a frequency approach was taken. For instance experts were asked :” If you have 10 patients, and you know that your staffing level is adequate , on how many of your patients assistive devices to reduce the risk of pressure ulcer will be used?”.

From what experts provided, the following probabilities (Table 4-3) for usage of assistive devices given all possible states of staff adequacy were determined.

Expert	Prob. Assistive Devices <i>Used</i> Given Staff Adequacy Situation	
	Adequate Staffing	In-adequate Staffing
Expert 1	100%	20%
Expert 2	100%	25%
Expert 3	90%	40%
Expert 4	90%	30%
Expert 5	> 95%	20%
Expert 6	90%	25%

Table 4-3. Expert opinion; probability of using assistive devices, given staff adequacy

To aggregate the experts' estimates, we will use the Bayesian methods for treating non-homogenous data, detailed in section 4.3.3.2.1.

Figure 4-10, shows the average distribution of the population variability distribution set for “Probability of Assistive Devices NOT Used, Given Adequate Staffing”, with mean 0.0477 and variance 0.0148.

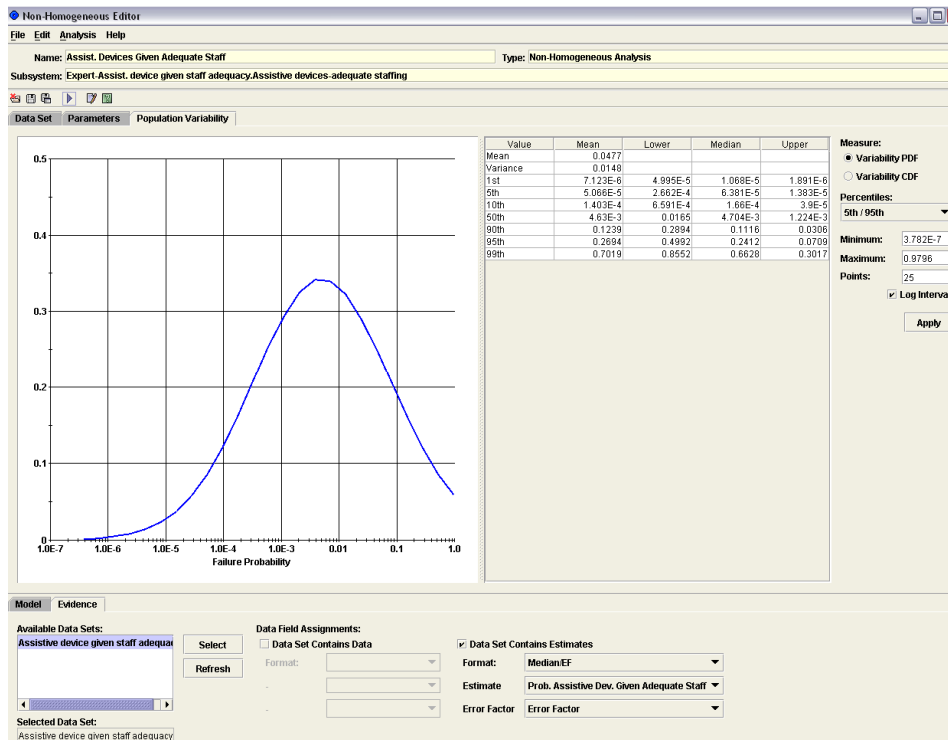


Figure 4-10. Average distribution of the population variability distribution set for “probability of assistive devices NOT used, given adequate staffing”

Also the same procedure estimates the “Probability of Assistive Devices NOT Used, Given Inadequate Staffing”. Figure 4-11, shows the average distribution of the population variability distribution set, with mean 0.7 and standard deviation 0.06.

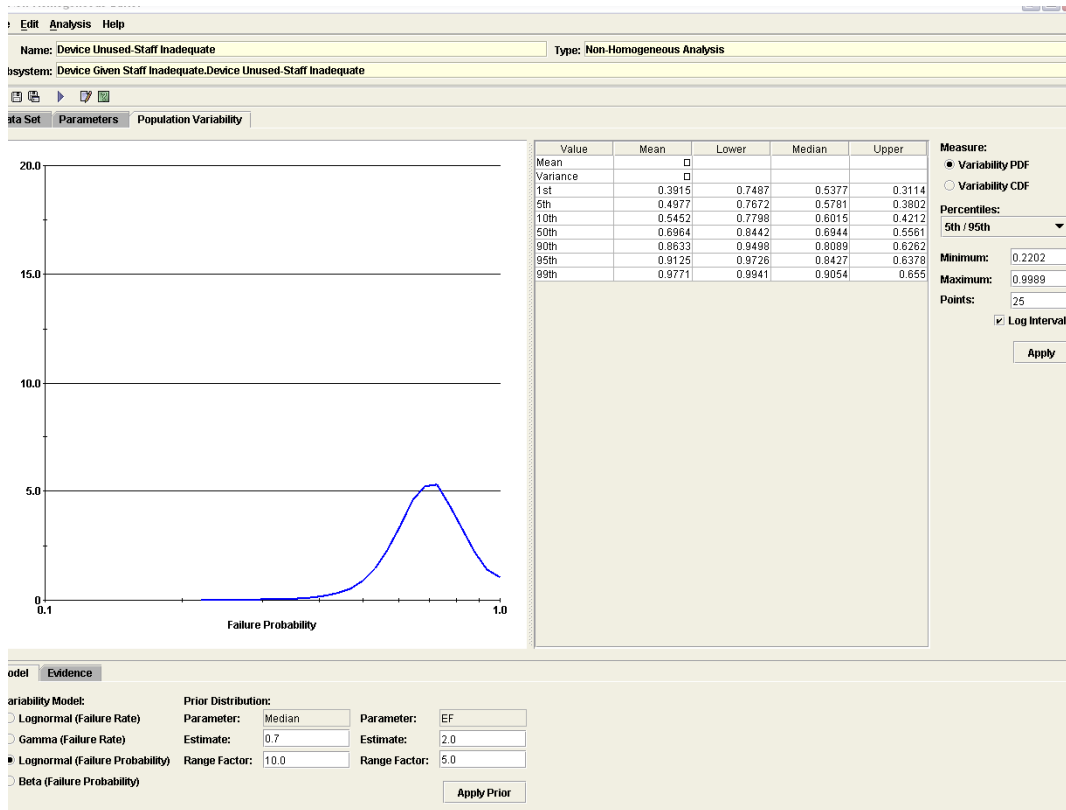


Figure 4-11. Average distribution of the population variability distribution set for “probability of assistive devices NOT used, given inadequate staffing”

8. Frequency of Move

States:

- i. Adequate = 1
- ii. Inadequate = 0

Inadequate frequency of move can increase the risk of developing pressure ulcers. As discussed before the frequency of move is influenced by many factors including Mobility, CNS impairment, Morbid Obesity, C-I Move and Staff Adequacy. As it was the

case for the nodes “Staff Adequacy” and “Assistive Devices”, no actual field data is recorded or available from data. To determine the conditional probabilities of adequate frequency of move and inadequate frequency of move given the different states of the abovementioned five parent nodes (cause factors), experts’ opinion is sought.

To facilitate the process of ease the process of expert elicitation we discussed and employed “parent divorcing technique” in 4.1.3.1.1. To implement this technique we created a dummy node that would capture the aggregate effects of four of five influencing factors of Frequency of Move and creatively called it “Aggregate Effect on Frequency of Move” with two possible states High and Low. The primary purpose of using this technique was to reduce the amount of information and estimates of probabilities we needed to elicit from experts, to ensure the estimates that are provided by experts are more robust and reliable. Using this dummy node, instead of original 5, Frequency of Move had now 2 parent nodes which reduces the number of question to be asked from experts from 32 (2^5) to 4 (2^2) questions. After explaining to the experts what the aggregate effect on frequency of move being high or low meant, following type questions were asked to obtain their judgment on the probability of frequency of move being adequate or inadequate given the states of its parent nodes (cause factors). For instance;

What is the probability of frequency of move being adequate, given that the aggregate effect on frequency of move is high and we DO have adequate staffing?

i.e.: $p(\text{Frequency of Move} = 1 | \text{Aggregate Effect} = 1, \text{Staff Adequacy} = 1)$

The following estimates in Table 4-4 were provided by the experts:

		Prob. Frequency of Move <i>Adequate</i> . Given the States of Aggregate Effect and Staff Adequacy			
Aggregate Effect		High		Low	
Staff Adequacy		Adequate	In-Adequate	Adequate	In-Adequate
Experts	Expert 1	95%	85%	99%	99%
	Expert 2	> 80%	50%	>95%	>90%
	Expert 3	70-80%	50-60%	>90%	85%
	Expert 4	>80%	40-50%	1	>90%
	Expert 5	98%	80%	> 99%	>90%
	Expert 6	90%	75%	1	>95%

Table 4-4. Expert opinion; conditional probability of adequacy of frequency of move

We aggregate these probabilities that have been provided by experts using the Bayesian method discussed in 4.3.3.2.1. The results of experts' estimates aggregation are as follows.

$$\Pr(\text{Frequency of Move} = \text{Inadequate} \mid \text{Aggregate Effect} = \text{High}, \text{Staff Adequacy} = \text{Adequate}) = \text{Log - Normal}(\text{mean} = 0.126, \text{variance} = 0.012)$$

Figure 4-12, shows the average distribution of the population variability distribution set.

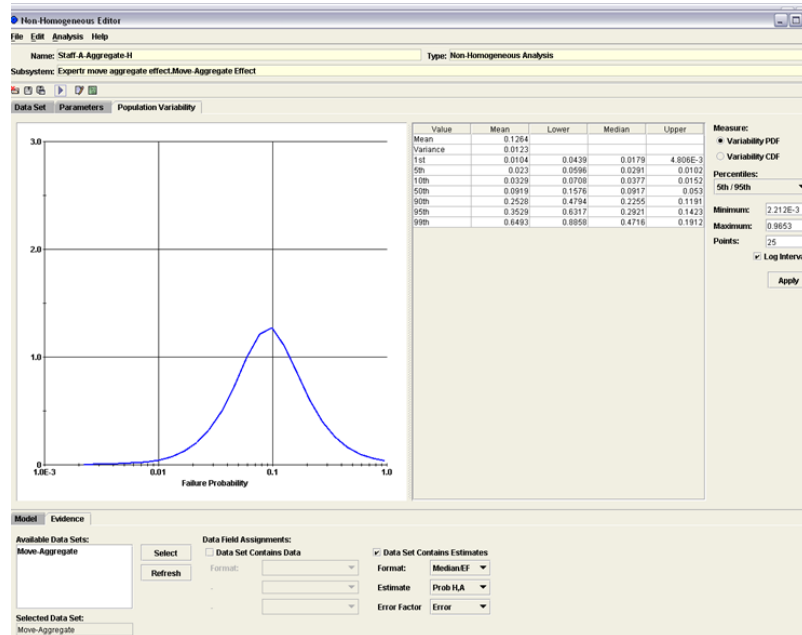


Figure 4-12. Average distribution of the population variability distribution set for probability of inadequate frequency of move (a)

$$\Pr(\text{Frequency of Move} = \text{Inadequate} \mid \text{Aggregate Effect} = \text{High}, \text{Staff Adequacy} = \text{inadequate}) = \text{Log} - \text{Normal}(\text{mean} = 0.336, \text{variance} = 0.0174)$$

Figure 4-13, shows the average distribution of the population variability distribution set.

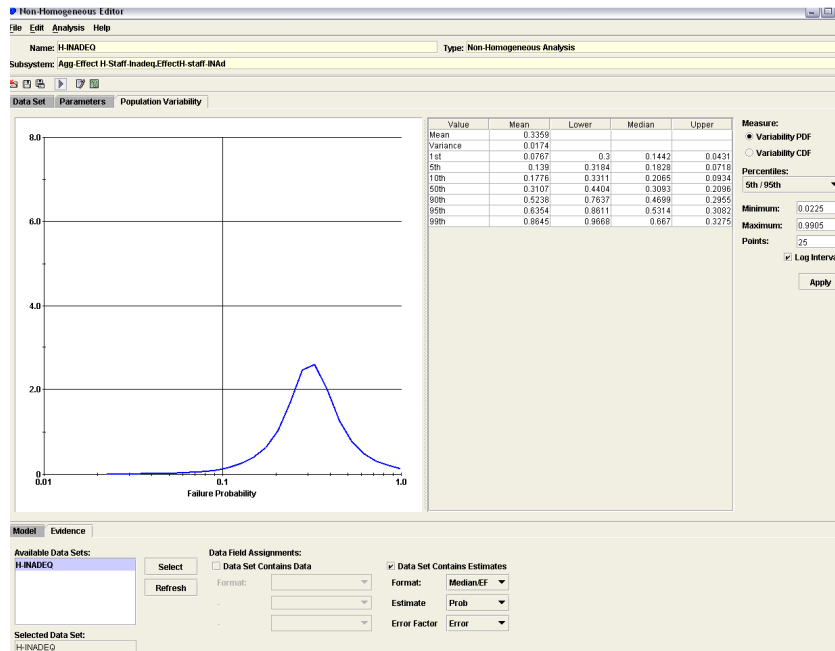


Figure 4-13. Average distribution of the population variability distribution set for probability of inadequate frequency of move (b)

$$\Pr(\text{Frequency of Move} = \text{In} - \text{Adequate} \mid \text{Aggregate Effect} = \text{Low}, \text{Staff Adequacy} = \text{Adequate}) = \text{Log} - \text{Normal}(\text{mean} = 0.025, \text{variance} = 0.006)$$

Figure 4-14, shows the average distribution of the population variability distribution set.

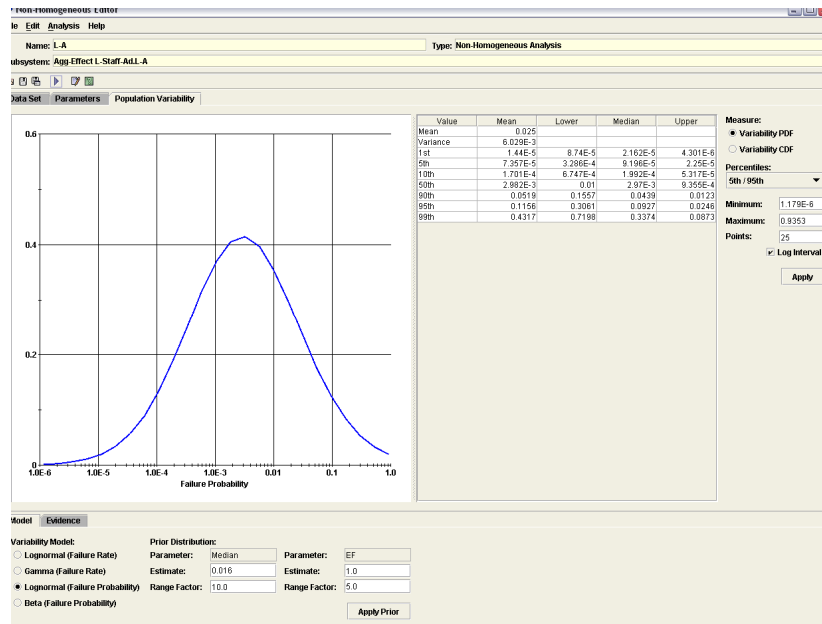


Figure 4-14. Average distribution of the population variability distribution set for probability of inadequate frequency of move (c)

$$\Pr(\text{Frequency of Move} = \text{Inadequate} \mid \text{Aggregate Effect} = \text{Low}, \text{Staff Adequacy} = \text{Inadequate}) = \text{Log - Normal}(\text{mean} = 0.06, \text{variance} = 0.0032)$$

Figure 4-15, shows the average distribution of the population variability distribution set.

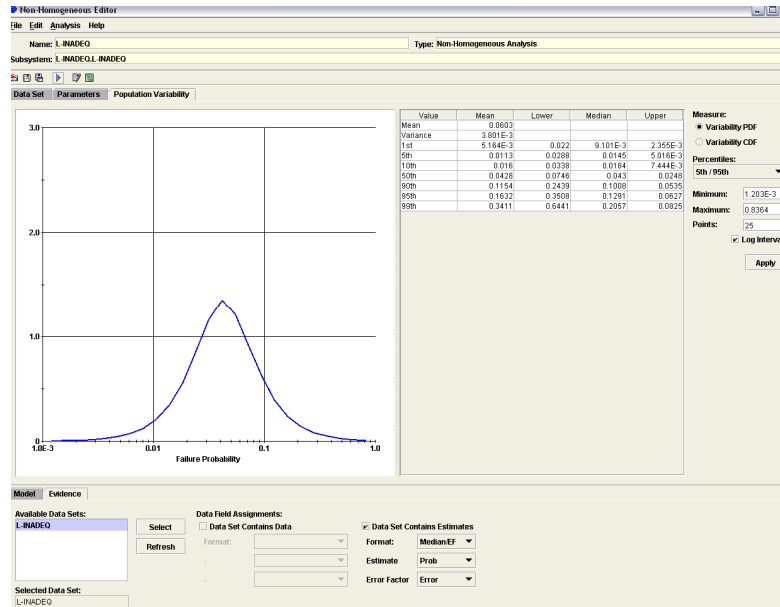


Figure 4-15. Average distribution of the population variability distribution set for probability of inadequate frequency of move (d)

The last step in quantifying pressure ulcer BBN, is to construct the conditional probability table for the node “Risk of pressure ulcer”, given all the risk factors. Given no hard data is available for “Assistive devices” and “Frequency of move”, we won’t be able to populate the CPT for “Risk of pressure ulcer” from data. On the other hand, if we want to elicit expert opinion for the CPT, it will be $2^5 = 32$ probability estimates, given that it has 5 parent nodes. Further, we have data for 3 of the parent nodes for 70,090 patients, that if we will leave rather unused if we only rely on subjective data. To make the best use of the existing hard data and to minimize the amount of information elicited from experts to ensure the reliability of the outcome of the elicitation, we will use the Noisy-OR gate algorithm, explained in section 5.1.3.3.1.2.

We have the probability of experiencing “Pressure Ulcer”, due to “Circulation impairment”, “Skin integrity”, and “Sensory impairment”, independently, from data. We

elicited experts' assessment on the probability that “Frequency of move”, and “Assistive devices” will independently cause pressure ulcer. Using these 5 probabilities (3 from data and 2 from experts), and using Noisy-OR gate algorithm, we can construct the whole CPT for “Risk of pressure ulcer”.

To elicit the probability that inadequacy of “Frequency of move”, and “Assistive devices” not being used, independent of any other factor will cause pressure ulcer, the experts were asked the following questions.

“To your opinion, out of 100 patients, how many are likely to experience pressure ulcer, because their frequency of move has been inadequate, regardless of any other risk factor?”

“To your opinion, out of 100 patients, how many are likely to experience pressure ulcer, because their no assistive device was used during their hospitalization, regardless of any other risk factor?”

Using Bayesian Framework that we discussed previously, we aggregate expert responses and assessments recorded in Table 4-5, and Table 4-6.

Expert	Probability of PU Due to Inadequate Fqcy. of Move
Expert 1	<10%
Expert 2	<10% and >5%
Expert 3	5%
Expert 4	15%
Expert 5	5%
Expert 6	< 5%

Expert	Probability of PU Due lack of Assistive Device Use
Expert 1	< 1%
Expert 2	2-3 %
Expert 3	1%
Expert 4	< 1%
Expert 5	5%
Expert 6	0

Table 4-5.Expert opinion; probability of pressure ulcer due to inadequate frequency of move (left)
 Table 4-6.Expert opinion; probability of pressure ulcer due to lack of assistive devices (right)

As a result the probability that inadequate frequency of move will cause pressure ulcer independent of other factors has a lognormal distribution with $\mu_{f_{qcy. on PU}} = 0.0672, \sigma_{f_{qcy. on PU}} = 0.038$

Similarly, the probability that not using assistive devices will cause pressure ulcer independent of other factors, has a lognormal distribution with parameters $\mu_{assist. devices on PU} = 0.018, \sigma_{assist. devices on PU} = 0.060$.

Figure 4-16 and Figure 4-17 show the average distribution of the population variability distribution sets.

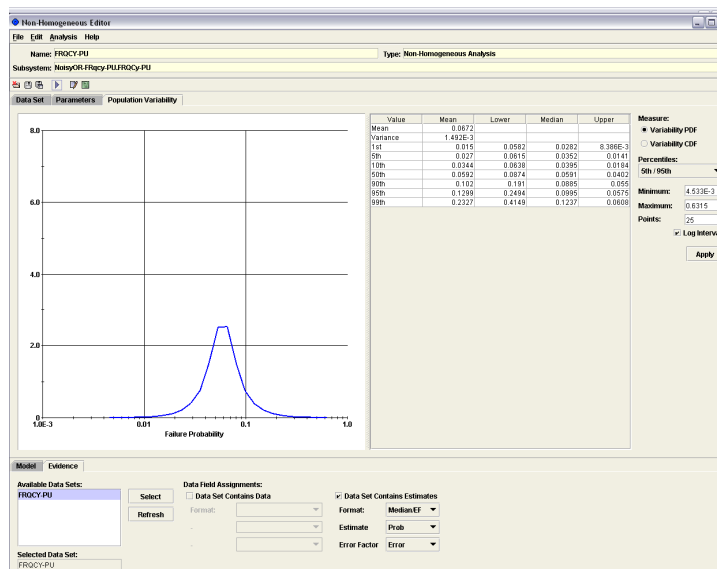


Figure 4-16. Average distribution of the population variability distribution for effect of frequency of movement on pressure ulcer

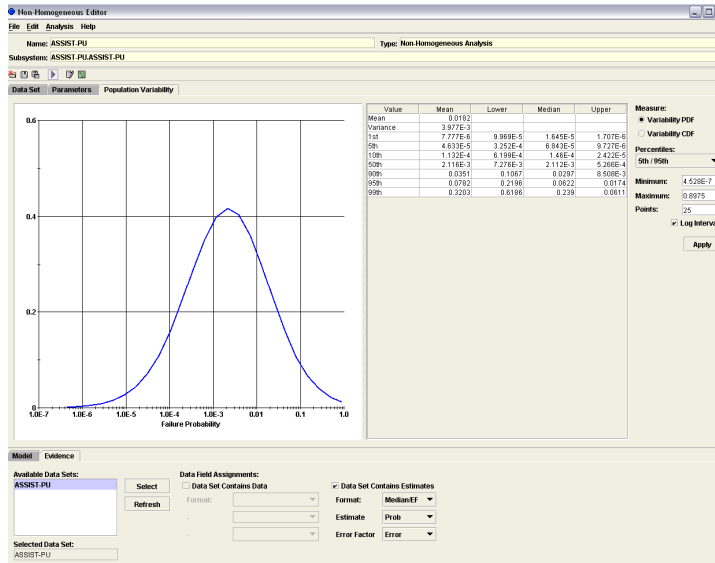


Figure 4-17. Average distribution of the population variability distribution for effect of using assistive devices on pressure ulcer

At this point, we have used GeNIe BBN software’s Noisy-Max option to quantify the BBN with Noisy-OR Gate procedure. As explained above, the probability of effect of each risk factor, on the risk of pressure ulcer (regardless of other factors), has been calculated from available data and experts’ opinion and is presented in Table 4-7.

Noisy-OR Gate				Total Records		70090
		PU	NPU	Total	Probability	Complement
Sensory=1, Circul=0, Skin=0		2	1299	1301	0.001537279	0.998462721
Sensory=0, Circul=1, Skin=0		11	8234	8245	0.001334142	0.998665858
Sensory=0, Circul=0, Skin=1		39	10684	10723	0.003637042	0.996362958
Frgc on PU (adequate)					0.0672	0.9328
Assistive Devic. On PU					0.018	0.982

Table 4-7. Probabilities of the effect of risk factors, independently, on pressure ulcer

Providing this input to GeNIe, we can calculate baseline probability of hospital acquired pressure ulcer (Figure 4-18). The model projects 3.3 E-3, for probability of pressure ulcer.

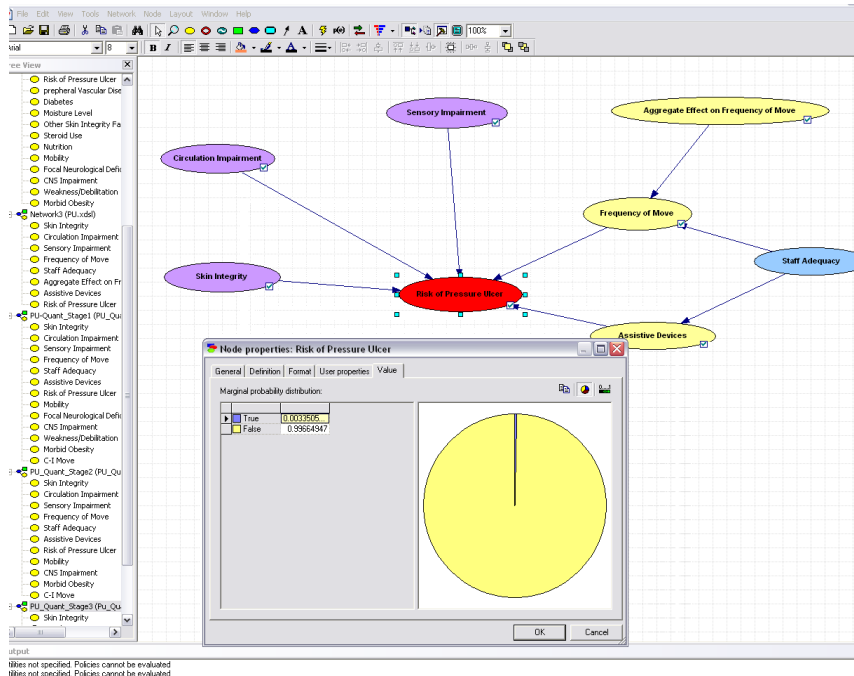


Figure 4-18. Probability of hospital acquired pressure ulcer

4.4 Pressure Ulcer BBN Validation

As discussed in section 3.6, validation and verification is a vital step in any type of model development in general. In developing the Bayesian belief networks for this study, we started with a basic draft of a model that contained the important factors and relations between the factors discussed in the literature and the input of one of the experts. We then consulted the domain experts extensively through multiple sessions of face-to-face interviews and reached to the consensus model that is presented here as the final version. This consensus was reached after many iterations to the point that all experts agreed that model is now presenting all the known major factors affecting the risk of pressure ulcer (and the risk of line infection in the case of vascular catheter associated infection). Naturally, peer review has been a crucial step in developing and qualitatively validating these models. In such a peer review of the BBN models,

some steps and methods, suggested by Marcot, et al. (2006) have been generally followed. These steps include:

- Introduction to BBN models
 - Introducing the concepts and general structure of the BBNs
 - Explaining how BBNs could be used to depict the causal and logical influences of key risk factors in pressure ulcer (and in line infection)
 - Explaining the general concepts of marginal (unconditional) probabilities for parent nodes and conditional probabilities of the child nodes
- Introduction and display of the specific BBN for pressure ulcer (and line infection) to review (this step is repeated after each iteration of the model based on previous interviews with the experts)
 - Explaining the objectives of the pressure ulcer (line infection) BBN models: to assess the stochastic effects of the physiological, intervention related, and hospital level factors on the risk of pressure ulcer (line infection)
 - Explaining the nodes in the model, what has been other experts rational to include or exclude a node, and also the linkage between the nodes
- Discussing the preliminary results
 - At the later stages of interview when the consensus on the factors and relations in the model is reached, with the available data the conditional probability table is constructed and a preliminary run of the model is presented. Also the concept of setting evidence and making inference is displayed. This specially helps and familiarizes the experts when they are asked for their opinion on some the probabilities that cannot be obtained from data

A form was designed based on these steps to guide the BBN development/ validation interviews and is available in appendices A&B.

4.4.1 Qualitative Validation of Pressure Ulcer BBN

Qualitative validation and verification of the Bayesian models in this research, is really built in the model development process. The example given below is the case of Pressure Ulcer BBN. The first draft of the model went through much iteration in expert interviews. Last, we asked our panel of experts to evaluate the last version of the model (the qualitative model) in following categories; model completeness, model accuracy, ease of understanding and perceived predictive validity, to ensure sufficient confidence in the structure of the model before proceeding to model quantification. This evaluation was performed through following question:

1. *Completeness.* From your perspective, to what extent does this model capture all important and relevant phenomena for the risk of pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that does not include some important and relevant details, whereas 100 would correspond to a model that includes all details that you consider important. What number would you assign?
2. *Accuracy:* From your perspective, how accurately or realistically does the model depict important factors that influence risk of experiencing pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that is unrealistic, over-idealized or inaccurate, whereas 100 would correspond to a model that is realistic and accurate. What number would you assign?

3. *Ease of understanding*: From your perspective, how easy is it to understand the overall logic of the model. On a scale from 0 to 100, 0 would correspond to a model that is difficult to follow, even with extensive explanation, and a 100 would correspond to a model that is readily understandable. What number would you assign?

4. *Perceived predictive validity*: From your perspective, if you were to use this model, how well could you predict the risk of pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that does not help at all with predicting effects, and a 100 would correspond to a model that predicts the effects very well. What number would you assign?

The expert's response to these questions, are summarized in Table 4-8.

	Completeness	Accuracy	Ease of Understanding	Predictive Validity
Expert 1	90	85-90	80	85-90
Expert 2	>90	>90	>90	>90
Expert 3	85-90	90	95	90
Expert 4	90-95	95	95	95
Expert 5	>90	>90	>95	70
Expert 6	>90	>90	>90	>90
Expert 7	95	95	95	85
Expert 8	90	90	90	90

Table 4-8. Expert opinion; qualitative evaluation of pressure ulcer BBN

All the experts unanimously agreed that to their opinion the pressure ulcer BBN and the line infection BBNs contained a comprehensive list of causing factors and the causing relations were accurately identified, but they felt more comfortable to score the models in the above four categories less than a perfect 100 because no model is ever perfect and there maybe factors (even though with marginal effects) that they have missed. This

builds a lot of confidence, at least in the qualitative representation of the models. The next section reviews the quantification challenges of Bayesian networks.

4.4.2 Quantitative Validation of Pressure Ulcer BBN

Relative frequency of hospital acquired pressure ulcer based on data for years 2003 to 2011 are recorded in Table 4-9 .

Fiscal Year	Hospital Acquired Pressure Ulcer	Total Admissions	Probability of Hospital Acquired Pressure Ulcer
2003	59	32616	0.00181
2004	67	33202	0.00202
2005	78	33351	0.00234
2006	76	33507	0.00227
2007	86	35195	0.00244
2008	104	36912	0.00282
2009	95	35769	0.00266
2010	53	36848	0.00144
2011*	22	13373	0.00165

Table 4-9. Relative frequency of hospital acquired pressure ulcer; 2003-2011

Compiling the Bayesian belief network, the model projected 0.0033, probability of pressure ulcer.

Available hospital acquired pressure ulcer data, reflected in table above, fits a Normal distribution (Figure 4-19) with mean 0.0022.

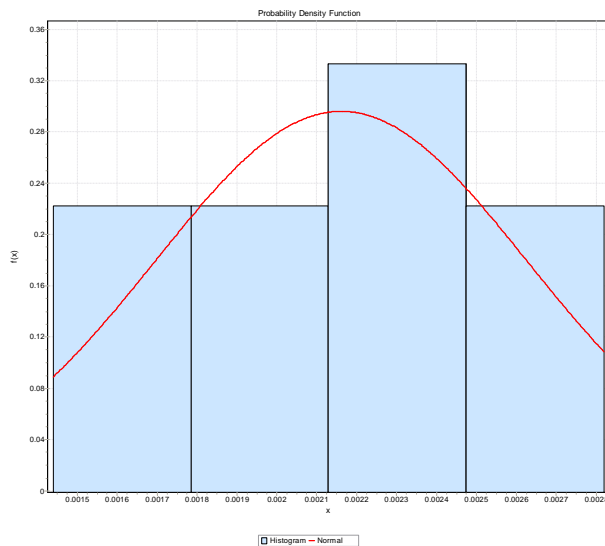


Figure 4-19. Distribution of the probability of hospital acquired pressure ulcer

This indicates that there is approximately 30% error in the prediction of the model, for pressure ulcer probability. This error is expected, chiefly due to the following reasons:

- a) The modifications we made to the model, without which the quantification of the model would have been impossible due to the absence of data
- b) For some of the nodes (e.g. Frequency of Move) we had to elicit expert opinion, and since no recorded data is available for such a variable we had no way of calibrating experts' inputs with the actual data.
- c) The quantification of model parameters is based on data for years 2008 and 2009, where we had reliable data available to us.
- d) Records only indicated whether the pressure ulcer was actually occurred while the patient was hospitalized (i.e. the patient was not admitted with pressure ulcer already present), since 2007. Prior to 2007 we only have the total number of patients with pressure ulcer (whether they acquired it in the hospital or not), and we had the hospital acquired pressure ulcer extrapolated.

For the above reasons, the projection of the model has a larger error, compare to the line infection BBN, where we had data available for all the variables for 2002-2009 (5% error). As more reliable data becomes available, one will be able to update the model with new information and obtain more precise results. This brings about the concept of model uncertainty, discussed in the next section. We have 2 ways of treating this model uncertainty; at the sub model level (BBN level), or at the hybrid model level (feeding a distribution as an input to the system dynamics model rather

than a point estimate). We have chosen the former, to avoid the propagation of the error at the BBN level to the hybrid model.

4.4.2.1 Treating Pressure Ulcer BBN Model Uncertainty

Predictive models are generally tools by which the modeler expresses his or her understanding of a particular unknown of interest. Since our knowledge about the true nature of this unknown is always incomplete, our expression of it inevitably involves uncertainty. In uncertainty analysis, we seek to address this lack of knowledge, with some confidence, in terms of the smallest range of possible values, which brackets the true value of the unknown of interest [Droguett, 1999]. These uncertainties are either associated with the values assumed by the model (“parameter uncertainty”) or with the structure of the model (“model uncertainty”). In this section our focus is on treating the model uncertainty in a Bayesian framework in order to improve the predictions made by pressure ulcer BBN model.

Predictions made by models contain an error; the error being the difference between the values produced by the model and the actual realization of the unknown of interest. In our context, $\mathcal{E} = p_t - p_m$, where p_t is the “true value” of pressure ulcer probability and p_m is the model’s prediction. In this case, an estimate for the true value of probability of hospital acquired pressure ulcer is the actual relative frequency of pressure ulcer for a particular year.

In order to improve the pressure ulcer probability prediction, we employ the Bayesian framework developed by Mosleh and Droguett (2008) to treat model uncertainty.

We are interested in assessing p_t , the true probability of pressure ulcer. Let's represent the BBN model's prediction as our evidence, $E = \{\underline{p}^*\}$. Our objective is to develop an uncertainty distribution of pressure ulcer probability, p , given the available evidence from our predictive BBN model. In its most general form, when we consider the pressure ulcer BBN model as a source of information, this uncertainty can be obtained as follows:

$$\pi(p|E) = \pi(p|\underline{p}^*) = \frac{L(\underline{p}^*|p)\pi_0(p)}{\int_p L(\underline{p}^*|p)\pi_0(p)dp}$$

where $\pi(p|E)$ is the posterior distribution of pressure ulcer probability p , $\pi_0(p)$ is the prior distribution of p , and $L(\underline{p}^*|p)$, the likelihood function, or the probability of observing evidence \underline{p}^* when the true value for probability of pressure ulcer p .

In this case, the form of information about the model is the performance of the BBN in predicting pressure ulcer probability. This information can be represented by the pair $(p_i^{Estimate}, p_i^{Actual})$ for year "i" for our model.

The relationship between the prediction of the model and the unknown p is given through the additive error model, where the model estimate is the true value of the unknown plus a random error $p_{m,i} = p_{t,i} + \varepsilon_i$, and $i = 1 \dots n$ represents the year of available performance data. Furthermore, we assume that performance data comes from a homogenous population. Since we are adopting an additive error model for pressure ulcer probability predictions, a flexible and practical form for the likelihood function is a

Normal distribution. Since we are dealing with the notion of probability that only assumes values between 0 and 1, we will use a truncated Normal distribution, bounded to 0 and 1. The likelihood function is therefore a truncated Normal distribution with mean $\overline{p_{PU_BBN}^{Estimate}} = p + b_{PU_BBN}$, where $b_{PU_BBN} = \bar{\varepsilon}$ is a bias factor and σ_{PU_BBN} , is the standard deviation:

$$L(p_{PU_BBN}^{Estimate} | p, b_{PU_BBN}, \sigma_{PU_BBN}) = \frac{1}{C\sigma_{PU_BBN}} \left(\frac{1}{\sqrt{2\pi}\sigma_{PU_BBN}} e^{-\frac{1}{2}\left(\frac{p_{PU_BBN}^{Estimate} - (p + b_{PU_BBN})}{\sigma_{PU_BBN}}\right)^2} \right)$$

where;

$$C = \Phi\left(\frac{1 - \overline{p_{PU_BBN}^{Estimate}}}{\sigma_{PU_BBN}}\right) - \Phi\left(\frac{0 - \overline{p_{PU_BBN}^{Estimate}}}{\sigma_{PU_BBN}}\right)$$

The posterior function of parameters $b_{PU_BBN}, \sigma_{PU_BBN}$ is:

$$\pi(b_{PU_BBN}, \sigma_{PU_BBN} | \varepsilon_1, \dots, \varepsilon_n) = \frac{L(\varepsilon_1, \dots, \varepsilon_n | b_{PU_BBN}, \sigma_{PU_BBN}) \pi_0(b_{PU_BBN}, \sigma_{PU_BBN})}{\int_{b_{PU_BBN}} \int_{\sigma_{PU_BBN}} L(\varepsilon_1, \dots, \varepsilon_n | b_{PU_BBN}, \sigma_{PU_BBN}) \pi_0(b_{PU_BBN}, \sigma_{PU_BBN}) db_{PU_BBN} d\sigma_{PU_BBN}}$$

where $\pi_0(b_{PU_BBN}, \sigma_{PU_BBN})$ is the prior distribution of $b_{PU_BBN}, \sigma_{PU_BBN}$.

For the case where the error terms depict a random behavior and display no trend, we can assume that $\varepsilon_1, \dots, \varepsilon_n$ are independent realizations of random variable ε , so we will have

$$L(\varepsilon_1, \dots, \varepsilon_n | b_{PU_BBN}, \sigma_{PU_BBN}) = \prod_{i=1}^n L(\varepsilon_i | b_{PU_BBN}, \sigma_{PU_BBN})$$

Substituting this into the posterior function of parameters $b_{PU_BBN}, \sigma_{PU_BBN}$, we will have

$$\pi(b_{PU_BBN}, \sigma_{PU_BBN} | \varepsilon_1, \dots, \varepsilon_n) = \frac{1}{k_1} \prod_{i=1}^n \frac{1}{\sigma_{PU_BBN}} e^{-\frac{1}{2} \left(\frac{\varepsilon_i - b_{PU_BBN}}{\sigma_{PU_BBN}} \right)^2} \pi_0(b_{PU_BBN}, \sigma_{PU_BBN})$$

k_1 , is a normalizing factor.

Finally, the likelihood of a new prediction of the model (for a future year) is

$$L(p_{PU_BBN}^{Estimate} | p, \text{ performance data}) = \int_{b_{PU_BBN}} \int_{\sigma_{PU_BBN}} \frac{1}{\sqrt{2\pi}\sigma_{PU_BBN}} e^{-\frac{1}{2} \left(\frac{p_{PU_BBN}^{Estimate} - (p + b_{PU_BBN})}{\sigma_{PU_BBN}} \right)^2} \frac{1}{k_1} \prod_{i=1}^n \frac{1}{\sigma_{PU_BBN}} e^{-\frac{1}{2} \left(\frac{\varepsilon_i - b_{PU_BBN}}{\sigma_{PU_BBN}} \right)^2} \pi_0(b_{PU_BBN}, \sigma_{PU_BBN}) d\sigma_{PU_BBN} db_{PU_BBN}$$

The corresponding posterior of the new model prediction is:

$$\pi(p | p_{PU_BBN}^{Estimate}, \varepsilon_1 \dots \varepsilon_n) = \frac{1}{k_2} L(p_{PU_BBN}^{Estimate}, \varepsilon_1 \dots \varepsilon_n | p) \pi_0(p)$$

Where k_2 , is a normalizing factor.

The model prediction for baseline probability of hospital acquired pressure ulcer is available from the pressure ulcer BBN; the actual probabilities of hospital acquired pressure ulcer are available in Table 4-9. Using this information, we can update and improve our BBN model predictions.

Figure 4-20 shows the distribution of the posterior function of the prediction of the Bayesian method for hospital acquired pressure ulcer probability, taking into the account

the performance of the BBN model. This posterior has a mean of 2.4 E-3, which compared to the average of actual data has 8% of error (compare to 33% error from the original prediction of the BBN).

All calculations have been done using “The Model Uncertainty Software,” a code developed by the Center for Risk and Reliability at the University of Maryland, College Park, in 2006 and it is available through this center.

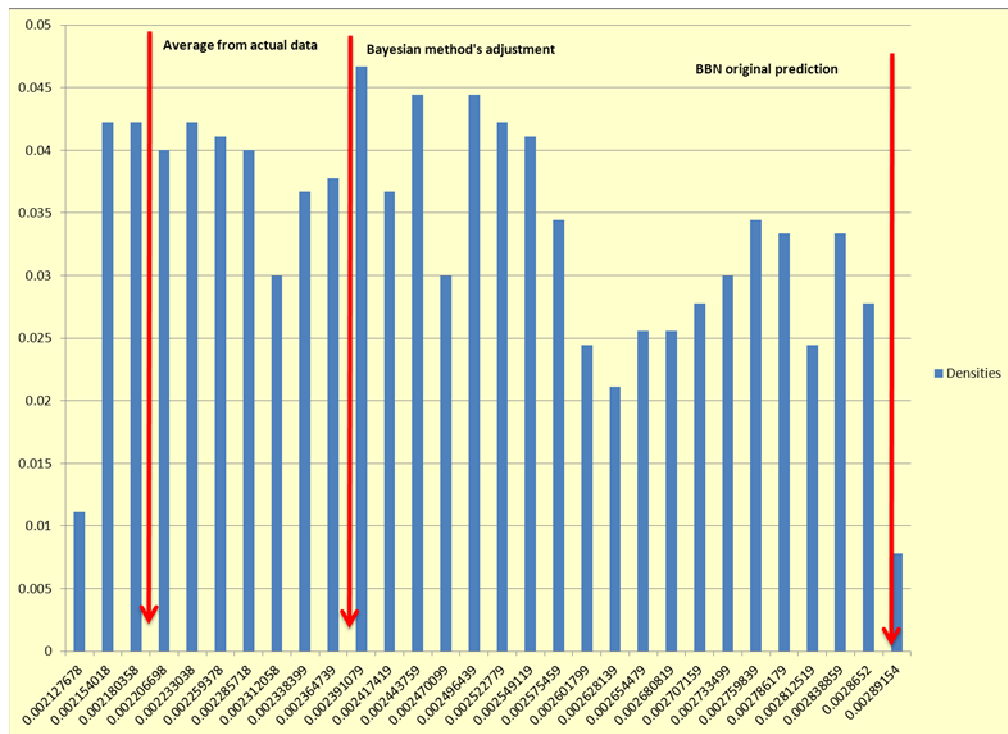


Figure 4-20. Posterior Distribution of Probability of Pressure Ulcer; BBN prediction adjusted using Bayesian model uncertainty method

5 Model Development for Adverse Events; Vascular Catheter-Associated Infection

In this chapter we will discuss the BBN model developed for risk of vascular catheter associated infection (i.e. line infection). The chapter provides a background on line infection, and development, quantification and validation of the line infection BBN.

5.1 Introduction

Central Venous Catheter (also called CVC, central line, or Vascular Access Device (VAD)), is a catheter that is placed into a large vein in the neck (internal jugular vein), chest (subclavian vein), or groin (femoral vein) to give medicines, fluids, nutrients or blood products to the patients. Intravascular catheters, as essential components of modern medical care, are one of the most commonly inserted medical devices in the United States, and also the most common cause of hospital acquired bloodstream infection, alongside urinary catheters. Unfortunately, most hospital acquired infections, in an already venerable patient population, are caused by the very same devices that are designed and used to provide lifesaving care. A study on medical intensive care units in the US has shown that 87% of bloodstream infections are attributed to central line (Trautner and Darouiche, 2004).

Vascular catheters, disrupt the protective barrier of the skin, and can potentially provide microorganisms with direct access to the bloodstream, which can cause local or systematic complications and in most extreme cases may cause death. In this section, we have developed a risk model, using Bayesian Belief Network formalism, to assess the

risk of experiencing vascular catheter infection, as a function of patient, and patient-provider factors.

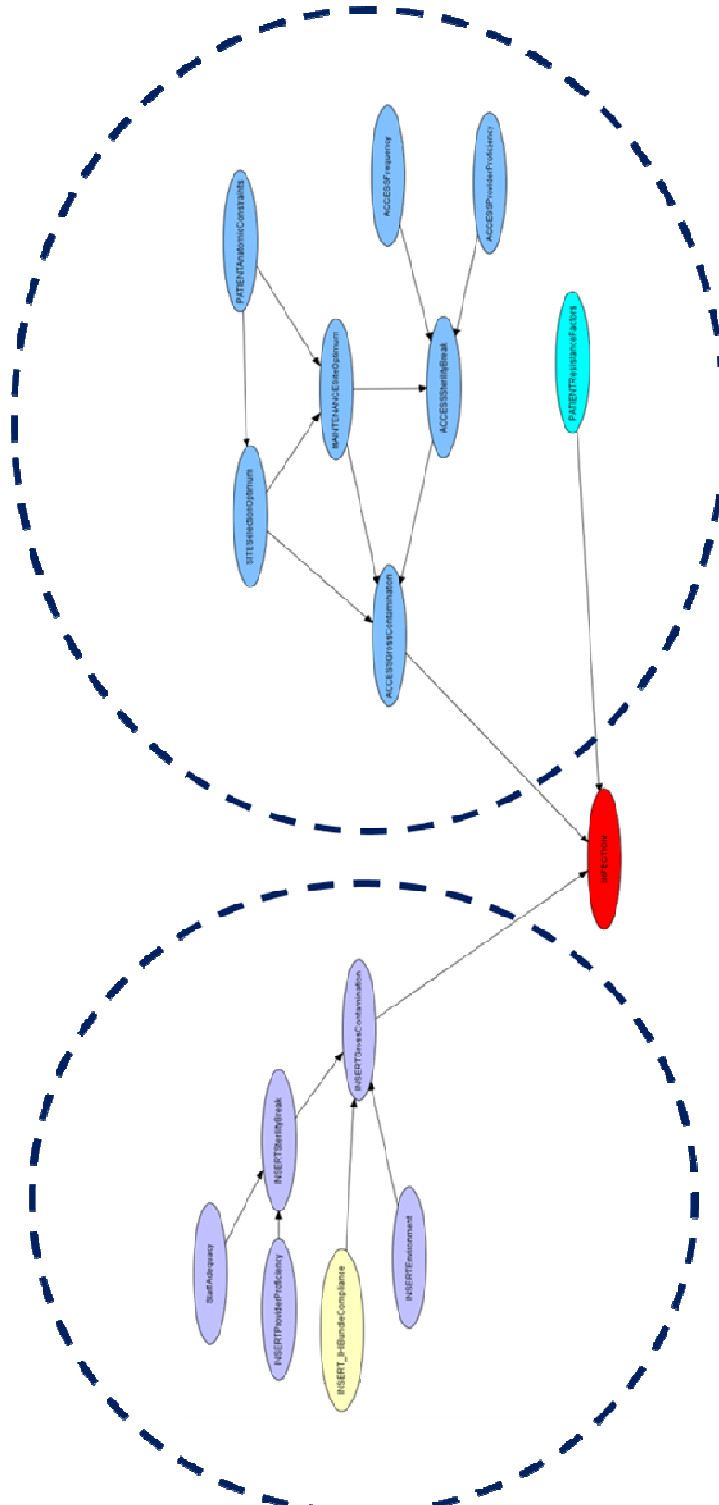
5.2 Vascular Catheter-Associated Infection BBN Development

A comprehensive literature review has been conducted to extract what researchers believe to be risk factors in line infection.

Richet et.al (1990), consider underlying disease, method of insertion, type of cannula (tube), type of dressing used, duration and purpose of catheterization as important risk factor, indicating that the impact of factors such as site of insertion, receipt of antimicrobial agents before, during and after catheterization, and the frequency of intravenous therapy (IV) are unclear. Moro et al. (1994), conclude from their study, that duration of catheterization, jugular insertion, transparent dressing, TPN (total parenteral nutrition), second catheterization period and skin colonization and hub colonization show significant association with catheter infection. In another study, Mahieu et al. (2001), find that catheterization duration, exit site colonization, hub colonization, insertion at bedside, whether or not patient is on antibiotics at insertion and TPN duration among important factors that may affect the risk of line infection.

A Bayesian Belief Network, that includes or reflects the factors introduced in literature as factors influencing risk of line infection, has been developed. Additionally factors that the panel of experts thought to be of importance have also been included. Figure 5-1 depicts this BBN. The validation process of the model has been detailed in section 5.4.

Figure 5-1. Vascular catheter related infection BBN



Insert Phase Risk Factors

Access, Use and Maintenance Phase Risk Factors

The risk factors in this BBN, are divided in two broad categories, Insert Phase Risk Factors, and Access, Use and Maintenance Phase Risk Factors.

A. Insert Phase Risk Factors:

This category of risk factors is concerned with issues and situations that may lead to contamination of the insert, and cause the micro-organisms to gain entry during the insertion procedure, and subsequently cause blood-stream infection. These factors include:

- Staff Adequacy: Availability of assistance to provider performing the procedure, during the insertion. Whether the unit has adequate staff available to perform the insertion helps reduce the chances of sterility break in the insert, and may also subconsciously reduce the chance of non-compliance of an individual, in following the safety protocols.
- Insert Provider Proficiency: provider's proficiency, experience and judgment during insertion phase influences the likelihood of insert sterility break. Also provider's proficiency decreases the likelihood of unsuccessful attempts to insertion and hence the probability of insert sterility breaks.
- Insert Sterility Break: Concerns unrecognized break in sterile technique.
- Insert IHI Bundle Compliance: The degree of compliance with the line insertion components of Institute for Healthcare Improvement's (IHI) bundle protocol (www.ihl.org), such as hand hygiene, skin preparation and use of barrier precautions.

- Insert Environment: Optimum environment for procedure, for example procedure room controlled environment versus bedside. This factor also encodes whether the line has been inserted in an emergency situation which increases the likelihood of insert contamination.
- Insert gross Contamination: Unrecognized gross contamination event.

B. Access, Use and Maintenance Phase Risk Factors:

This category of risk factors is concerned with issues and situations that may lead to contamination of the access area, while the line is in place and infection may be introduced to the blood stream during the maintenance phase of the line. These factors include:

- Patient Anatomic Constraint: That influence;
 - i. Site selection: for instance subclavian versus less desirable jugular or femoral vein
 - ii. Choice of de novo insertion versus less desired change over guide wire
 - iii. Need to perform site maintenance procedures for example dressing change
- Site Selection Optimum: Addresses the anatomic setting of catheter:
 - i. Subclavian vein site: Inserting the line in the chest area
 - ii. Jugular vein site: Inserting the line in the neck area
 - iii. Femoral vein site: Inserting the line in the groin
- Maintenance Site Optimum: Optimum maintenance of the insertion site; the integrity, manipulation and the state of dressing

- Access Frequency: Frequency of port access which influences the likelihood of access sterility break. In other words the more frequently the port is accessed, the higher the exposure of the access site to contamination would be.
- Access Provider Proficiency: Provider’s proficiency, experience and judgment during the access or maintenance procedures affect the likelihood of access sterility break.
- Access Sterility Break: Unrecognized break in sterility during the port access or use of the device.
- Access Gross Contamination: Gross contamination of the site, i.e. access port or actual skin insertion site
- Patient Resistance Factor: Physiological and pharmacological factors, influence resistance and susceptibility to infection.
- Infection: Determines the probability of blood-stream infection given all possible states of the risk factors.

5.3 Vascular Catheter-Associated Infection BBN Quantification

To carry out the quantification of the line infection BBN, certain modifications had to be made to the structure of the BBN without compromising the integrity and accuracy of the model. In the consensus model shown in Figure 5-1, “Staff Adequacy” and “Insert Provider Proficiency” influence an intermediate node “Insert Sterility Break”, which in combination with “Insert IHI Bundle Compliance” and “Insert Environment” , affect the probability of “Insert Gross Contamination”. A similar node “Access Gross

Contamination” also exists in the maintenance phase of catheter lines. Gross contamination of insert and gross contamination of access port; influence the probability of bloodstream infection.

Truth is, the contamination event is rarely witnessed but physicians know that it had to have occurred given the influencing factors. In the databases used to extract data for line infection BBN quantification, no record was found with documentation of a gross contamination. If that was ever obvious to the clinician, they removed the line immediately and started over. What we have recorded data on, are the precursors in the causal structure. Since data was available on whether these influencing factors (Staff Adequacy, Insert Provider Proficiency, Insert IHI Bundle Compliance, and Insert Environment) were present for each patient, and we also knew whether this patient had an infection or not, we could directly calculate the effect of these factors on the risk of infection, eliminating the intermediate nodes without compromising the accuracy or the integrity of the model. The same is true for the access and maintenance phase of the model, and we can safely remove the intermediate node of “Access Gross Contamination”, and directly measure the strength of the effect of the influencing factors (maintenance phase risk factors) on the risk of infection. Modifying the BBN, for the purpose of quantification, based on the justification provided above, results in the line infection BBN, depicted in Figure 5-2.

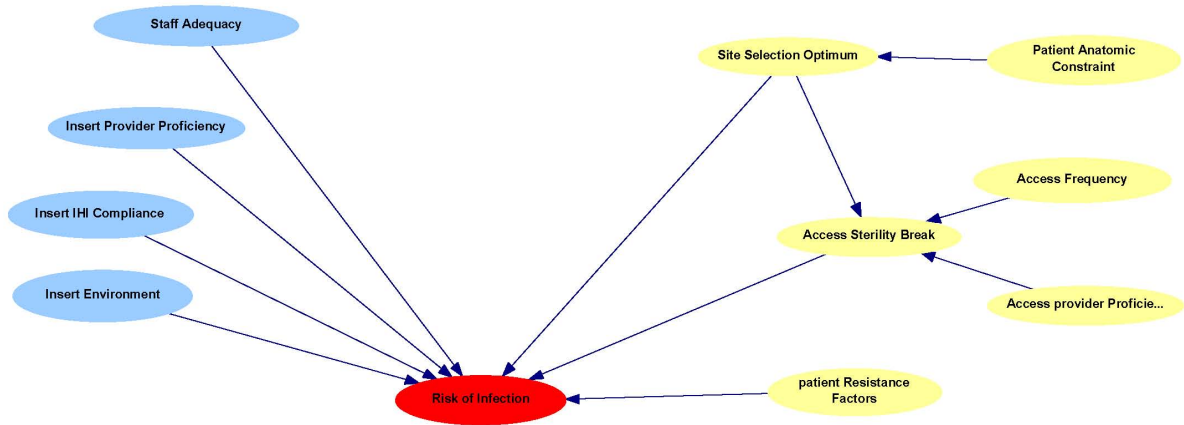


Figure 5-2. Line infection BNN; transformed

Moreover, we have extracted line infection data from ICU patients, as the data were most reliable and the results could be extrapolated to the entire hospital. In any given institution most of the lines are in the ICU and very few lines on the floors, and in fact some institutions have rules where you cannot have a line on the floors.

We have extracted and analyzed 12897, ICU patient records from October 2001 to September 2009. Figure 5-3 shows a few records of the data.

EventID	Occurrence Month	Blood Stream Infection (0=N;1=Y)	InsertEnvironment (0=Unknown;1=Optimal;2=Suboptimal)	InsertProviderProficiency (0=Unknown;1=Expert;2=Novice)	StaffAdequate (0=Unknown;1=Y;2=N)	Bundle Compliance - (0=unknown;1=full compliance;2=partial compliance)	PtAnatomicConstraint (0=Unknown;1=Y;2=N)	Site Selection (0=Unknown;1=Optimal;2=Suboptimal)	AccessFrequency (0=Unknown;1=High Frequency; 2=Low Frequency)	AccessProviderProficiency (0=Unknown;1=Expert;2=Novice)	PatientResistanceFactors (0=Unknown;1=High Resistance Capability;2=Diminished Resistance)	AccessSterilityBreak (0=Unknown;1=Rare/Minor;2=Common/Major)
1	Oct-01	1	2	1	0	0	0	2	1	1	1	2
2	Oct-01	1	2	2	2	0	2	2	1	1	1	2
3	Oct-01	1	1	2	1	0	2	1	1	2	2	1
4	Oct-01	1	2	2	1	0	2	1	2	2	2	0
5	Oct-01	0	1	2	1	2	2	2	1	1	1	1
6	Oct-01	0	2	2	1	1	2	1	1	1	1	2
7	Oct-01	0	1	1	1	2	2	2	1	1	1	1
8	Oct-01	0	1	1	1	2	1	1	1	1	2	1
9	Oct-01	0	1	2	1	0	2	2	1	1	2	1
10	Oct-01	0	1	2	1	0	2	1	2	1	1	1
11	Oct-01	0	1	2	1	0	2	2	2	1	2	1
12	Oct-01	0	1	2	1	0	2	2	2	1	1	1
13	Oct-01	0	2	1	1	1	0	2	1	1	1	1
14	Oct-01	0	1	1	1	1	2	2	2	1	2	1
15	Oct-01	0	1	2	1	1	2	2	1	1	2	1
16	Oct-01	0	2	2	1	1	2	1	1	1	1	1
17	Oct-01	0	2	1	1	1	2	1	1	1	1	1
18	Oct-01	0	1	1	1	1	1	1	2	1	2	1
19	Oct-01	0	1	2	1	2	2	2	2	2	1	1
20	Oct-01	0	2	2	1	1	2	1	2	1	1	2
21	Oct-01	0	1	2	1	1	2	1	1	2	1	1
22	Oct-01	0	2	1	1	1	2	2	2	1	1	1
23	Oct-01	0	1	1	1	1	2	2	1	2	1	1
24	Oct-01	0	1	2	1	1	2	2	2	2	1	1
25	Oct-01	0	2	2	1	1	2	1	1	1	2	1
26	Oct-01	0	1	1	1	1	2	2	1	2	1	1
27	Oct-01	0	2	1	2	1	1	2	1	2	1	1
28	Oct-01	0	2	2	1	1	2	1	2	1	2	1
29	Oct-01	0	1	1	1	1	2	1	2	2	2	1

Figure 5-3. Sample data records used for line infection BBN quantification

To calculate the marginal and conditional probabilities needed for quantification of the BBN, we have used parameter learning option of GeNIe BBN software developed by University of Pittsburg, PA. To learn parameters and populate the conditional probability table for an existing network with defined structure, after importing both the network and the data (in the form of a text file), we need to create a mapping between the variables defined in the network and variables defined in the data set. Data records that are unavailable could be identified as “N/A”, when importing data to the software. The results of the calculated marginal and conditional probabilities are discussed below.

1. Staff Adequacy

States:

- i. Adequate: 0.84
- ii. Inadequate: 0.16

If we could confirm through documentation of a nursing note or the bundle checklist that a staff member was available (for assistance), it was declared N/A. If there was evidence of staffing, but also evidence of other distracting or competing activities, we also declared N/A. Note that the probability calculated here is used as a prior, and will be updated with the system dynamics model's output on staff adequacy.

2. Insert provider Proficiency

States

- i. Expert : 0.59
- ii. Novice: 0.41

3. Insert IHI Bundle Compliance

States

- i. Full Compliance : 0.84
- ii. Partial Compliance: 0.16

Data is extracted straightforward based on the bundle elements. Notably, some of the factors for partial compliance are potentially weak influencers of infection, but we did not sub segment the compliance.

4. Insert Environment

States

- i. Optimal : 0.85
- ii. Suboptimal: 0.15

When there was evidence in the electronic record documenting where the procedure took place, we were able to determine whether the insert environment was optimal. An optimal environment indicates that the environment was the ICU patient room (a semi-controlled environment) or the operating or procedure room. A common example of a score of suboptimal environment, would be the trauma room, ED or regular non-ICU clinical unit, or during an emergency resuscitation.

5. Patient Anatomic Constraint

States

- i. True: 0.11

ii. False: 0.89

6. Site Selection Optimum

States

i. Optimal: 0.63

ii. Suboptimal: 0.37

7. Access Frequency

States

i. High frequency: 0.62

ii. Low Frequency:0.38

This is probably the most difficult to score. We based it on the concurrent use of drugs and invasive physiological measurements that were carried out. There is no widely accepted standard, but if the patient had 4 or more IV infusions, and concurrent central venous monitoring, we scored high frequency. All else were low frequency. If the patient died quickly, or there was poor documentation of route of administration of drugs or use of the line, it was declared N/A.

8. Access Provider Proficiency

States

i. Expert: 0.78

ii. Novice: 0.22

This is also difficult to measure, but we assessed based on the primary nurse that was assigned. Patients actually have several nurses caring for them throughout a hospitalization, but we focused on their primary consistent nurse coverage.

9. Access Sterility Break

States

- i. Rare : 0.9
- ii. Common/Major : 0.1

If we found documentation of a break in the clinical annotation, we scored a major break. If there was documentation of access and documentation that the dressing was intact and that sterile procedures were followed for access and no documentation of a break, we gave them a rare break. If there was no documentation of access (i.e. absence of documentation that the site was inspected, dressing was intact, sterile procedures used, etc., then we were skeptical — i.e., suspected that there was a documentation problem, not necessarily the absence or presence of a break. So here, we declared N/A.

10. Patient Resistance Factors

States

- i. High Resistance Capability: 0.58
- ii. Diminished Resistance: 0.42

If the patient was profoundly immunosuppressed, as evidenced in a diagnosis like 'Stasis Post Bone Marrow Transplant', or 'Acute Lymphoma', then we readily scored them as diminished resistance. Some patients were receiving broad spectrum antibiotics, and if so, we scored them high resistance.

11. Risk of Infection (Bloodstream Infection):

States

- i. True
- ii. False

If a bloodstream infection was identified, and other sources were ruled out, we scored true. Note that if a patient developed a bloodstream infection, had a central line, but another source of infection was possible, the patient was scored as a false for line infection.

Relative frequency of line infection based on data for years 2002 and 2009 are recorded in Table 5-1. Relative frequency of hospital acquired bloodstream infection; 2002-2009.

Year	Probability of Line Infection
2002	0.050
2003	0.039
2004	0.037
2005	0.032
2006	0.023
2007	0.022
2008	0.024
2009	0.019

Table 5-1. Relative frequency of hospital acquired bloodstream infection; 2002-2009

Compiling the Bayesian belief network, using the above probabilities and the conditional probability table calculated, the probability of line infection produced by the model, using all data from 2002 to 2009, is 0.0322, which is very close to the relative frequency we can obtain from total number of line infections in these years, divided by total number of

cases ($\frac{389}{12897} = 0.0302$).

5.4 Vascular Catheter-Associated Infection BBN Validation

Both qualitative and quantitative validation of line infection BBN is discussed in this section.

5.4.1 Vascular Catheter-Associated Infection BBN Qualitative Validation

Similar process was employed to construct the Bayesian model for line infection, using the same panel of experts. After necessary changes and modifications were made to the first draft of the model, through multiple interviews with each expert, we asked experts to evaluate the qualitative model with respect to completeness, accuracy, ease of understanding to ensure sufficient confidence in the structure of the model. Table 5-2 contains the results of this evaluation.

	Completeness	Accuracy	Ease of Understanding	Predictive Validity
Expert 1	90	90	80	90
Expert 2	>95	>95	>90	>95
Expert 3	90	90	90	90
Expert 4	90-95	90-95	95	90-95
Expert 5	>95	>90	>95	>80
Expert 6	>90	>90	>90	>90
Expert 7	95	95	95	90
Expert 8	85	85	85	85

Table 5-2. Expert Opinion; qualitative evaluation of line infection BBN

5.4.2 Vascular Catheter-Associated Infection BBN Quantitative Validation

Compiling the Bayesian belief network, using the above probabilities and the conditional probability table calculated, the probability of line infection produced by the model, using all data from 2002 to 2009 (Table 5-1), is 0.0306.

Line infection probabilities for years 2002 to 2009, could be represented by a Weibull distribution ($\alpha = 3.2206, \beta = 0.03119$) (as shown in Figure 5-4 below), with mean 0.03125 (and 10 percentile and 90 percentile values of 0.019 and 0.049 respectively). The value produced by the BBN for the probability of line infection has 3% error compared to the mean of the distribution of line infection probability from data.

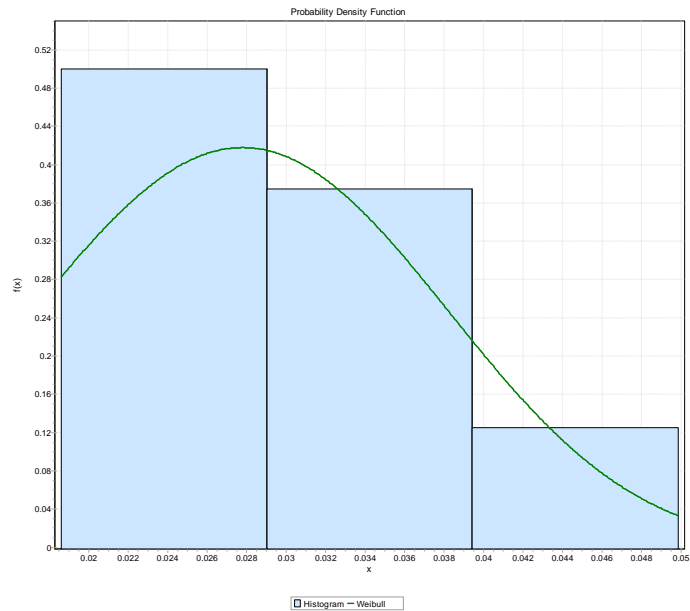


Figure 5-4. Distribution of line infection probability; 2002-2009

6 Dynamic Model of Hospital-level Factors Affecting the Risk of Adverse Events

Using system dynamics formalism, we are going to demonstrate how organizational level and policy level contributions to risk change over time, and how policies and decisions may affect the general system-level contribution to adverse event risk. The dynamic model developed in this study, also captures the feedback of organizational factors and decisions over time and the non-linearities in these feedback effects. Given a baseline level of certain adverse events risk that every patient is exposed to, due to his or her physiological conditions and caregiver's intervention, we are interested to see how this baseline risk may change because of the decisions/policies at the hospital level with regards to pressures to reduce operational costs, optimize length of stay and investments in proactive safety interventions. The baseline risk for specific adverse events that are of interest in this study (the list is given in section 3.4.1), which accounts for patient level and provider-patient level factors are the outputs of the BBNs. Two such BBNs have been developed in this study for the risk of pressure ulcer and the risk of line infection. The organizational level factors that may affect these baseline risks however, are addressed in the system dynamics model discussed in this chapter.

6.1 Dynamic Model Development

6.1.1 Introduction

The focus of the dynamic model is a specific hypothesis. Combination of increasing costs and decreasing reimbursement has created tremendous financial constraints for healthcare organizations and additionally, insurers have increased pressure by imposing penalties for adverse events (at least certain adverse events listed in 3.5.2). This situation leaves hospitals in the following risk-relevant positions; they will have few resources to invest proactively in safety and they will have to make operational decisions (such as reduction in staffing) that focus on reducing costs, which nonetheless may increase the risk.

The use of system dynamics formalism will help us model the evolution of internal/external financial and decision/policy factors on safety state of the organization. Our emphasis is on capturing the dynamic changes in safety state of the hospital as a function of reimbursement, financial penalties imposed by external agencies and productivity pressures. In other words we are modeling how, changes in the safety state of the organization, subsequently increase or decrease the risk of specific adverse events.

6.1.2 Model Developing Process

The process of developing the qualitative part of the model is very much similar to the development of the qualitative model for the Bayesian Networks. We started with a rough draft of the model, that represented how financial standing of the hospital leads to cost reduction strategies and constraints hospital ability to invest in safety, and how these decisions may increase risk of adverse events and the feedback effect of these adverse

events on hospital's finances. Peer review has been a vital step in developing and qualitatively validating the dynamic model, as well as the Bayesian models. Our general framework for building (and qualitatively validating) the qualitative model using domain experts, follows the steps suggested by Marcot et al. (2006) in peer review of Bayesian Networks, which are quite applicable and useful in any influence diagram type, qualitative causal model building. The process of building the qualitative part of our system dynamic model, through interviews with subject matter experts, follows these guidelines;

- *First round of the interviews*
- Introduction to system dynamics formalism/models.
- Explaining how system dynamics could be used to depict causal relationships, non-linearities in feedback effects and change over time.
- Explaining the hypothesis of interest in the study, that how the financial wellbeing of the hospital effects the decisions to reduce costs and constraints investments in safety, and how these decisions may affect the risk of adverse events and finally the feedback effect of these adverse events on hospital's finances. This step is conducted using the highest level of abstraction of the model.
- Explaining the detailed version of the first draft of the model that was developed based on literature and with the help of one of the experts.
- Asking the experts to review the first draft of the model and alter/modify it in any way they see fit, including adding, deleting, modifying any factor or causal relation/loop from this draft.

- *Second round of the interviews*
 - Discussing the second draft of the model, that contains all the modifications that experts had made to the first draft in the first round of interviews. Also explaining the justifications provided by other experts on the modifications they had possibly made to the first draft of the model.
 - Asking the experts to review the second draft and make any modifications necessary.
- *Third round of the interviews*
 - Discussing the final, consensus model containing all modifications (after a few iterations)
 - Asking experts to rate the qualitative model, from 0-100 , with respect to completeness, accuracy, ease of understanding and perceived predictability
 - Eliciting experts' quantitative assessment on some of the nodes of the model (this step is discussed in more detail in section 4.3.2)

6.1.3 The System Dynamics Model

6.1.3.1 Introduction

Decision making, policy and action relating to safety of the hospital is strongly influenced by its financial state. Complex relationship between operational expenses and reimbursement by external agencies influences overall revenue. Generally, increases in operational expenses and decreases in reimbursement rates lead to a revenue gap. U.S.

hospitals have consistent responses to such revenue gap. Usually one or a combination of the following strategies, are taken to deal with this revenue gap:

- Limiting costs associated with managing patients, primarily by reducing length of stay (LOS) to a minimum
- Reducing staffing
- Limiting expenditures and investments in proactive safety interventions

To demonstrate and the value and effectiveness of the hybrid technique used in this research (combination of system dynamics and Bayesian networks), we have developed a model that explores how risk of specific adverse events changes over time as a function of several system constraints. In particular, we are examining the impact of reimbursement, financial penalties and productivity pressures on the risk of hospital-acquired adverse events such as infections, medication errors, falls and other patient injuries. In detail, the model includes:

- The impact of increasingly (financially) unfavorable reimbursement policies that have been established by private and public insurance companies
- New financial penalties imposed by private and public insurance companies in response to specific adverse events (i.e., new policies under which reimbursement for care is not reimbursed when a hospital-acquired adverse event occurs)
- Intense production pressures and pressures to reduce length of stay (LOS) in order to reduce costs assure reimbursement by insurance companies.

The model captures the interactions of these factors on the probability of adverse events. By using system dynamics formalism, the model captures the effects of feedback

reinforcement on risk over time. The model captures the delayed effects of relaxing throughput pressure on the risk of adverse events, as reduced revenue eventually leaves a hospital with little or no resources to commit to proactive safety investments or maintain the current safety measures. In other words, the model not only considers organizational factors, but also takes into account the policy environment in which such decisions are made and how this changes over time.

6.1.3.2 Model Structure; Key Variables and Important Feedback Loops

The basic structure of the dynamic model to depict the organizational decisions and strategies to control the revenue gap and the effects of these strategies on the adverse event risk, has been built on a well-established and very common system dynamics concept, known as downward spiral or vicious circles. The concept has been widely used in modeling business processes, and social contexts. For instance, for a manufacturing or service providing company, unfavorable revenues may result in cost cutting strategies. Cost cutting strategies will cause service/quality loss which translates into revenue loss due to inferior service/quality which clearly worsens the unfavorable revenue situation that triggered this “vicious circle” or “downward spiral”. Figure 6-1 depicts this reinforcing loop.

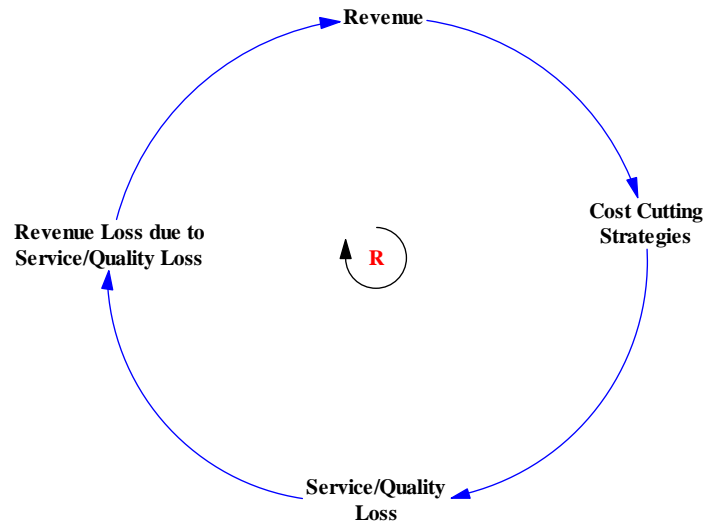


Figure 6-1. Revenue deficit-loss of service downward spiral

Figure 6-2, the highest level of abstraction of the model, depicts the key variables influencing the occurrence of adverse events, based on literature, interviews with clinical experts and field observations. The hypotheses that are to be validated with data, through this model are basically captured in the loop structures, explained in this section. These loops examine the effects of the strategies adopted by hospitals in response to their revenue gap on the risk of specific adverse events, and the feedback effect of these adverse events on the hospital's financial wellbeing.

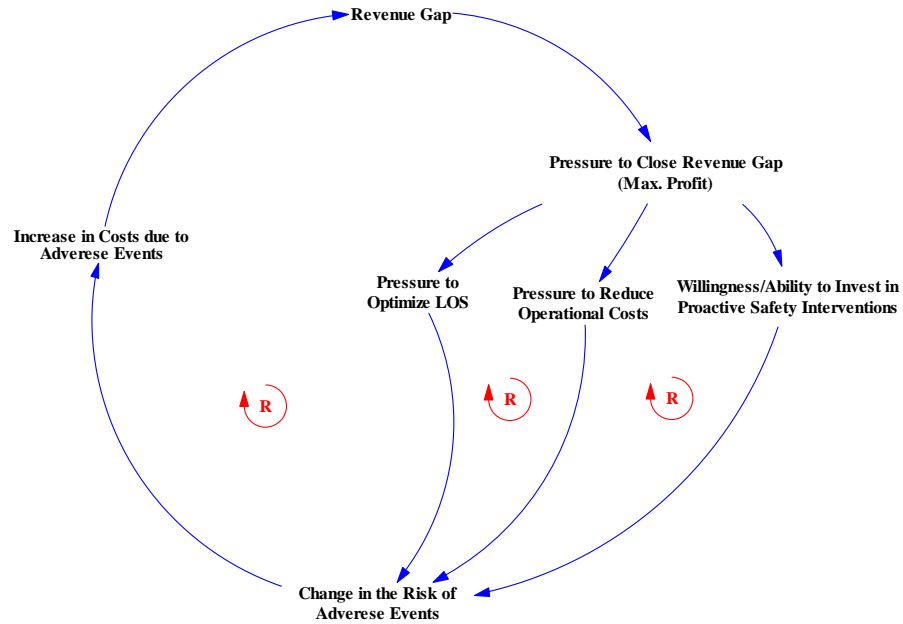


Figure 6-2. System dynamics module; highest level of abstraction

Hospital is providing a service and it costs a certain amount of money to deliver this service. Part or all of this cost will be reimbursed by patient's insurance company, according to a predefined arrangement. Bases on the average number of days from previous year that each patient has stayed in the hospital for a certain diagnosis-related group (DRG), the insurance company informs hospitals of the amount of money that the hospital will be reimbursed for treating patients with that specific DRG. This reimbursement is subject to denial on the part of the insurance companies, for a number of reasons. One such a reason that is still being debated is the possibility that the hospitals will not be reimbursed for the cost they bear to care for certain hospital acquired conditions. The livelihood of a hospital, or any service providing organization for that matter, depends on the profitability of that organization. This imposes a pressure on managers and other levels of decision making in the hospital to maximize the differential

cost/reimbursement, and in other words close the revenue gap. As a response to the pressure to close this gap, U.S. hospitals usually adopt one or a combination of the following strategies.

6.1.3.2.1 Strategy 1: Reducing Length of Stay

The unfavorable cost reimbursement differential in hospital creates pressure to maximize this differential (maximize profit). One of the ways this pressure manifests itself is through the pressure to reduce LOS to a minimum.

In this model we have considered that the risk of specific adverse events that are of interest in this study could be affected in four different ways:

- Change in the adverse event risk due to shortened LOS,
- Change in the adverse event risk due to prolonged LOS
- Change in the adverse event risk due to understaffing, and
- Change in the adverse event risk due to lack of investment in proactive safety interventions

The pressure to optimize LOS to the minimum required increases the chance that a patient will be discharged prior to readiness and before all needs are met. Hence the probability that the patient will experience an adverse event increases. Additionally, if an adverse event occurs to this patient, he or she will have to return to the hospital for treatment of the experienced AE, and therefore he or she is required to spend more time in the hospital, which makes the patient prone to the adverse events due to prolonged

LOS. A decrease in safety investments also affects the probability of adverse events in all three categories of AE.

6.1.3.2.2 Strategy 2: Reducing Staffing

Another strategy to respond to the pressure to reduce operational costs is reduction in the level of staffing. Staff reduction is based on a wishful thinking that it would be possible to care for the same number of patients with less staff, without degrading safety and quality of care. This in short term reduces costs and sets cost/reimbursement differential on a favorable path, but in long run it may increase the probability of AEs, and may lead to increases in cost in many ways.

6.1.3.2.3 Strategy 3: Reducing Proactive Safety Investments

While it is in hospital's best interest to invest proactively in safety, the unfavorable revenue may lead to policy decisions that avoid investment in proactive safety investments. This in the short time will save the hospital some money but in the long run increases the risk of experiencing adverse events. This strategy is analogous to the concept of reduced investment in maintenance in engineering systems.

6.1.3.2.4 Feedback Influences

On the other hand, increase in the number of adverse events increases the cost for hospital in many ways. It costs hospital more to provide care for the complications caused by adverse events. As a result of the AE, patients have to stay longer in the

hospital and the hospital will bear the associated costs, and there will be no reimbursement for the cost of treating the experienced AE. Additionally hospital's capacity will decrease and the hospital will be unable to admit new patients. Also more challenging reimbursement policies will be imposed on the hospital in future. On top of that, hospital will suffer the loss of trust and good will on patient's end.

Starting from the high level model in Figure 6-2 and after many revisions the model evolved to the more detailed version shown in Figure 6.3. Section 6.2 discusses the quantification of this model.

Figure 6-3. System dynamics module; detailed version

6.2 Dynamic Model's Quantification

Based on available clinical data and experts' assessment, the formulas for each of the nodes, the formulas representing the relationships between the nodes, and in case of the constants in the model the respective values of the nodes, have been derived and calculated. This section provides detailed discussion on these formulas and values.

Operating Margin

One of the measures hospitals use to evaluate how well they are doing financially, is "Operating Margin". Operating margin (OM) is the ratio of operating income divided by revenue, and operating income is simply their revenue minus cost.

$$\text{Operating Margin} = \frac{\text{Revenue} - \text{Cost}}{\text{Revenue}}$$

Hospitals' fiscal year starts on October first and ends on September thirty first. We have operating margin, operating dollars, and cost and revenue data available to us for years 2003 through 2011, where 2011 data is partial, from October first 2010 to January first 2011. Table 6-1 contains the available financial data.

Year	OM	OM Dollars	Revenue (R)	Cost (C)
2003	-3.20%	\$ (14,600,000.00)	\$456,250,000.00	\$470,850,000.00
2004	1.20%	\$ 13,600,000.00	\$1,133,333,333.33	\$1,119,733,333.33
2005	3.40%	\$ 33,700,000.00	\$991,176,470.59	\$957,476,470.59
2006	2.60%	\$ 27,500,000.00	\$1,057,692,307.69	\$1,030,192,307.69
2007	4.70%	\$ 52,388,770.15	\$1,114,654,684.10	\$1,062,265,913.94
2008	-0.90%	\$ (10,253,429.78)	\$1,139,269,975.04	\$1,149,523,404.81
2009	1.30%	\$ 14,847,535.95	\$1,142,118,149.97	\$1,127,270,614.02
2010	2.10%	\$ 24,044,442.35	\$1,144,973,445.35	\$1,120,929,003.00
2011	3.50%	\$ 40,174,255.76	\$1,147,835,878.96	\$1,107,661,623.20

Table 6-1. Fiscal data

6.2.1 Soft Factors in the Model: "Pressures"

To understand how the financial stress, leads to managerial decisions at the hospital level to reduce loss and maximize profitability, we have introduced the concept of "Pressure to Close Revenue Gap" into the model, which forces the decision makers to adopt cost cutting strategies that may influence the risk of specific adverse events. In other words, this pressure creates pressure to reduce length of stay, pressure to reduce operational costs (mainly staffing) and affects the organization's willingness and ability to invest proactively in safety.

Whilst these pressures are extremely real and visible to every health care professional in clinical settings (including all the experts interviewed for this study), they are soft, human-oriented factors that because they have not been reflected explicitly in any analysis or database for that matter, are very challenging to model mathematically. Naturally, expert judgment can play a significant role in, at least subjectively, formulating these concepts. To be able to elicit expert's opinion on the relationship between "Revenue Gap" and "Pressure to Close Revenue Gap", the concept of pressure has to be communicated with the experts in such a way that the outcome of the elicitation process is sufficiently reliable. In other words, we had to make sure that the experts had a clear understanding of what our vision was about these pressures.

In modeling soft factors of this kind, different approaches could be found in system dynamics literature. Some of these soft factors are easier to model than the others, mainly because some type of proxy (which we possibly have some data on) could be used in modeling them. For instance, Sterman (2000), in modeling generic structure for a labor

capacitated process uses “Desired Completion Rate / Standard Completion Rate” to formulate “Schedule Pressure”. On the other hand, there are also examples in the literature that in modeling some the soft factors, have relied solely on subjective assessment of the experts and experts’ belief about the certain entity being modeled and have not represented the soft factor using a measure that corresponds to a physical entity. For instance, McCabe (1998) has used a scale of 0-1 to model airline employees’ moral, and Cooke (2004) has used a similar scale to represent management and personnel commitment to safety in modeling the operational risk in mining industry.

In modeling pressures in this model, even though interviews with experts revealed that the notion is rather clear to clinical professionals, to further ensure the reliability of the elicited information from experts on the soft factors of the model we decided to find a measure for these factors that was more tangible. This would allow us to be more confident that the questions from experts about the form, shape, effect and value of these pressures is being communicated correctly and us analysts, as well as our experts understand and mean the same thing about these soft factors. For this we relied on the notion of *elasticity*, used frequently in economics.

Elasticity, in economics, is the ratio of the percent change in a variable to the percent change in another variable, and is used as a tool to measure the responsiveness of a function to changes in parameters in a dimensionless manner. For example “price elasticity of demand”, E_d ;

$$E_d = \frac{\text{percent change in quantity demanded}}{\text{percent change in price}} = \frac{\Delta Q_{\text{demanded}} / Q_{\text{demanded}}}{\Delta p / p}$$

Gives percentage of change in demand in response to one percent change in price, given all other factors such as income remain constant.

We will utilize the concept of elasticity, in its general sense, to elicit experts' opinion on the formulation of the four soft factors in the model (Pressure to Close Revenue Gap, Pressure to Optimize LOS, Pressure to Reduce Operational Costs, and Willingness/Ability to Invest in Proactive Safety Interventions). That is in interviewing experts for these soft factors, we anchored our questions on the elasticity concept so that the "Pressure to Close Revenue Gap", etc. is connected to an actual physical measure, but the outcome of the elicitation is still subjective.

1. Pressure to Close Revenue Gap

Financial stress forces decision makers to adopt cost cutting strategies. We have established in the qualitative model, that in U.S. hospitals these strategies mainly include optimizing LOS, reducing operational costs and limiting proactive expenditure on safety.

To model "Pressure to Close Revenue Gap" we asked the experts:

"Lower or negative operating margins will force management to take one or a combination of cost cutting strategies. The worse the operating margin gets the higher the chances of adopting these strategies. Given the graph below (Figure 6-4), please mark what ranges of operating margins corresponds to the likelihood of some kind of cost cutting decision being enforced, to your experience. Use the range 0-1 for pressure, where 1 corresponds to the maximum level of pressure"

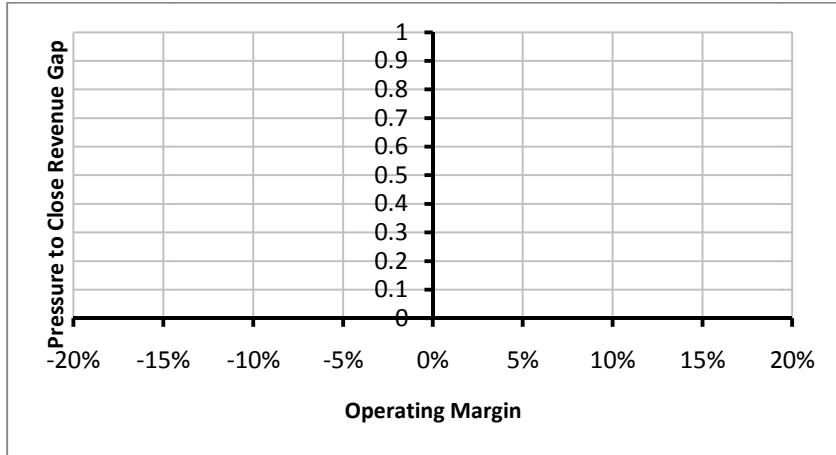
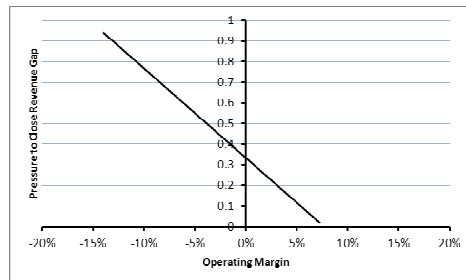
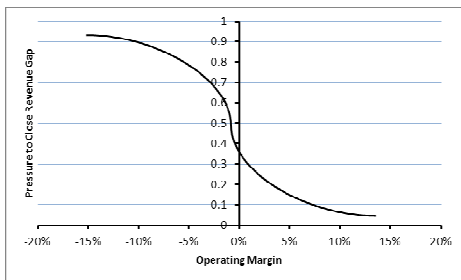


Figure 6-4. Operating margin-pressure to close revenue gap relationship

To facilitate the elicitation process, the experts were given a few forms and graphs and were asked to select the form they thought best represented the relationship between operating margin and pressure to close revenue gap. They were also asked to mark the threshold values on the graphs. Figure 6-5, shows the format options that were presented to the experts. The complete interview sheet could be found in appendix C.



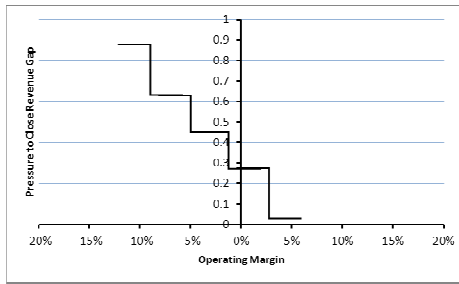


Figure 6-5. Operating margin-pressure to close revenue gap relationship; format options

Most experts chose the form of a step function, and a few experts picked the inverse S shaped function. None of the experts believed that this relationship could be represented with a linear format. We picked a step function form, and aggregating the turning points on the graphs provided by experts, we arrived at the graph depicted in Figure 6-6.

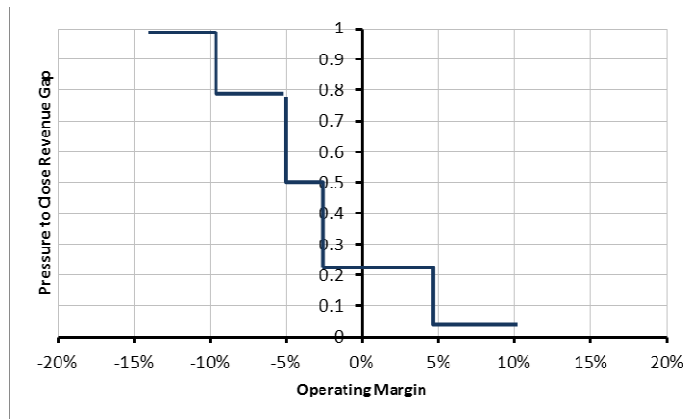


Figure 6-6. Expert opinion; operating margin/pressure to reduce revenue gap relationship

Given the outcome of the elicitation for the relationship between “Operating Margin” and “Pressure to Close Revenue gap”, we have to be able to manipulate this relationship in the system dynamics model, that is, the node “Pressure to Close Revenue Gap” is an intermediate node between “Operating margin” and ultimately “Risk of Specific Adverse Events” (actual field data is available for both of them) and we have to be able to

calibrate the “Pressure to Close Revenue gap” to get the best fit to the actual data. In order to do this though, we have to convert this step function to the closest parametric, functional forms.

The best parametric function that fitted the experts’ input is a negative Sigmoid function in the form of:

$$PCRG = [1 / (1 + \text{Exp}(p_2(RG + p_1)))] + p_3$$

Where:

PCRG: Pressure to Close Revenue Gap

RG: Revenue Gap

And, p_1, p_2, p_3 are the parameters of the model. Fitting the negative Sigmoid function to experts’ assessments reveals the optimum values for the parameters as;

$$p_1 = 0.059$$

$$p_2 = 75.57$$

$$p_3 = 0.044$$

Figure 6-7, shows the function for “Pressure to Close Revenue Gap” using the optimum values for the parameters.

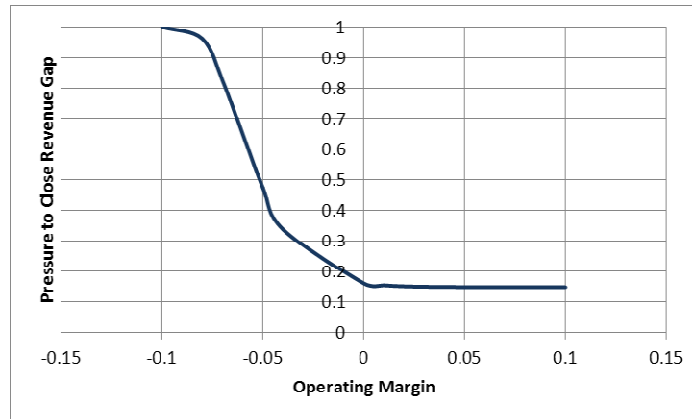


Figure 6-7. Operating margin/pressure to reduce revenue gap relationship; functional form

As it is evident from the graph, experts believe that organizational tolerance for loss is rather very small, the pressure to make decisions in order to close the revenue gap spikes rapidly and dramatically, when the hospital is experiencing financial distress. Additionally, there is always a pressure even though very small to maximize profit.

2. *Pressure to Optimize LOS*

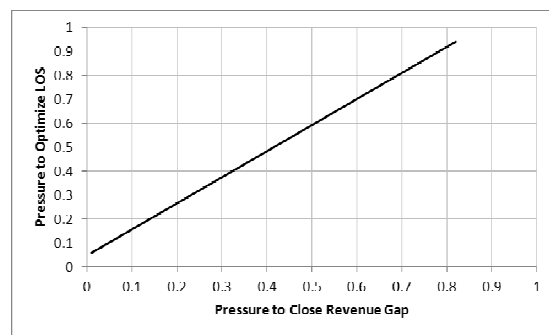
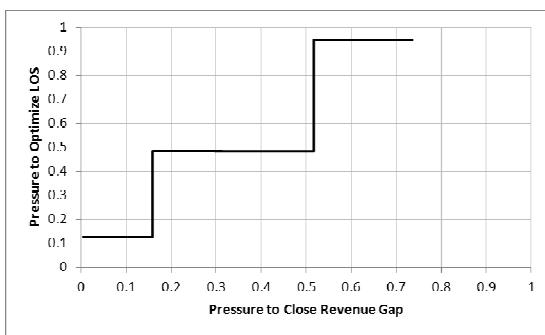
Pressure to close revenue gap, leads decision makers towards cost cutting strategies. One of those strategies is optimizing LOS or reducing LOS to the minimum required, which we have argued that affects the probability of experiencing certain adverse events. In formulating the relationship between pressure to close revenue gap, and pressure to optimize LOS, again we have utilized the concept of elasticity in eliciting experts' assessments of such relationship.

A tangible and real entity to clinical professionals, that we could relate to the pressure to optimize LOS, is the concept of “*utilization review meetings*”, where physicians systematically visit unit by unit and decide which patients could theoretically be discharged. These meetings are real actions that take place in part in response to the need

to reduce LOS, and they actually increase in frequency (maybe even twice daily) as the need to reduce LOS increases. Nevertheless, depending on the physical conditions of the patients these patients may or may not achieve actual discharges of the patients. To be able to reliably elicit experts' opinion on "Pressure to Optimize LOS" we will use the utilization review meetings' frequency as a proxy to evaluate the level of pressure to reduce LOS.

To model "Pressure to Optimize LOS" we asked the experts:

"Given your assessment on "Pressure to Close Revenue Gap", as a function of "Operating Margin", what is the likelihood of observing a change (increase or decrease) in frequency of the utilization review meetings given the level/ranges of organizational pressure to close revenue gap, maximize profitability and minimize costs?. Please mark these ranges in the graph that best represents this relationship to your opinion (Figure 6-8)"



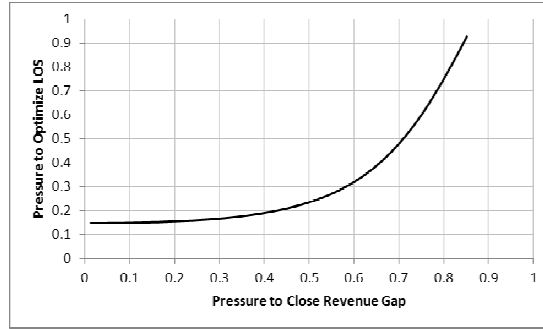


Figure 6-8. Pressure to close revenue gap-pressure to optimize LOS relationship; format options

Most experts chose the form of a step function, and a few experts picked the exponential function, with different slopes. None of the experts believed that this relationship could be represented with a linear format. We picked a step function form, and aggregating the turning points on the graphs provided by experts, we arrived at the graph depicted in Figure 6-9.

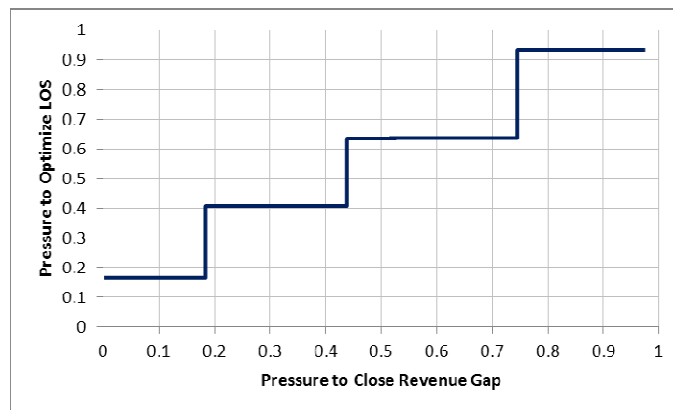


Figure 6-9. Expert opinion; pressure to reduce revenue gap relationship/pressure to optimize LOS relationship

Interestingly enough, most experts marked the intercept greater than 0, meaning that regardless of the financial pressures, there is always some level of pressure to optimize

LOS. In this case also, similar to the case of “Pressure to Close Revenue Gap”, and with the same rationale, we have to convert this step function to the closest parametric, functional forms. The best parametric function that fitted the experts’ input is a negative exponential function in the form of:

$$PRLOS = 1 - \text{Exp}(-p_4 \times PCRG) + p_5$$

Where:

PRLOS: Pressure to Reduce LOS

PCRG: Pressure to Close Revenue Gap

and , p_4, p_5 , are model parameters.

Fitting the negative exponential function to experts’ assessments reveals the optimum values for the parameters as;

$$p_4 = 1.55$$

$$p_5 = 0.149$$

Figure 6-10, shows the function for “Pressure to Optimize LOS” using the optimum values for the parameters.

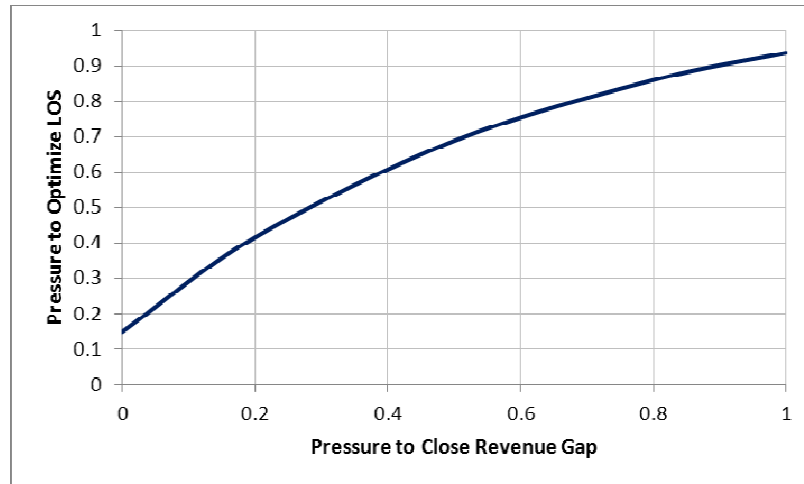


Figure 6-10. Pressure to reduce revenue gap relationship/pressure to optimize LOS relationship; functional form

3. *Pressure to Reduce Operational Costs*

Another strategy to close revenue gap and maximize profitability, is reducing the operational costs. The largest piece of operational costs in hospitals that is directly impacted by such strategies is admittedly staffing. Reduction in staffing that could potentially have great impact on the risk of certain adverse events, concerns reduction in nursing staff. Physicians are usually not affected greatly by decisions to reduce staff, since they are able to generate revenue. Before we talk about formulating this node of the model, we will briefly discuss how staffing is structured in the hospitals.

When it comes to staffing, ideally the decision maker knows that optimally, certain number of staff with certain level of expertise is needed to care for certain number of patients with a certain range of problems. So the decision maker looks at the pool of nurses and if enough nurses are not available they use what is known as *Per Diem* nurses, that are nurses who work on temporary assignments usually through specialized

placement agencies or through hospital staffing pools. So in an ideal situation, where there is no expenditure constraints you can always optimally staff using per diem nurses, but these services are very expensive and they cost much more than using staff nurses (i.e. the hourly rate).

There are a number of ways, in which hospitals take action to control and reduce staffing costs;

1. There are many different ways of staffing and each impacts the costs differently. Probably the most expensive way, is to staff with hospital's employees and filling the gaps with agency nurses. The first response to pressure to reduce staffing is eliminating agency or per diem nurses. Although there are of course exceptions, for instance if there is an epidemic of Influenza and the hospital has none of its own staff available, agency nurses have to be used regardless.
2. The second approach to reduce staffing costs is eliminating overtime within the hospital's staff. Overtime is very costly for the hospitals (e.g. time and a half in OR), and to assign nurses, they have to make sure to use nurses that are not overtime and almost mandate them to cover the shift the hospital needs. Hence in tight financial constraints, elimination of overtime is another measure that is taken to reduce staffing costs.
3. Depending on the type of unit you are staffing, there is some flexibility in terms of the composition of the staff. There are Registered Nurses (RNs, have completed college nursing program and have passed a national licensing exam), Licensed Practical Nurses (LPNs), Patient Care Technicians (PCTs, usually a high school degree with some additional training). They have all been trained in

patient care, but with different levels of education and experience. In a perfect situation, the person who is staffing a floor or a unit would use RNs. But that is an expensive model, so the tasks that could be done by LPNs, hence the composition changes to say, 2/3 LPNs and 1/3 RNs. As the next step down, there might be some tasks assigned to LPNs that PCTs can do as well, so you change some of the LPNs in your staffing composition and replace them with patient care technicians.

To summarize, eliminating per diem nurses, eliminating staff over time and changing the composition of nursing staff are tangible, physical actions that take place in response to financial constraints. These changes and cuts nevertheless are bounded by some regulation as well, and there are certain staffing ratios that have to be maintained (this ratio slightly varies from state to state), and although there may be instances where these regulatory ratios are not maintained, but most hospitals closely follow this regulation and avoid running the risk from licensing stand point.

Now to model “Pressure to Reduce Operational Costs”, given the 3 staffing cost reduction strategies discussed above, we asked the experts:

“Given your assessment on “Pressure to Close Revenue Gap”, as a function of “Operating Margin”, what is the likelihood of observing an action regarding the three strategies to reduce operational costs (staffing costs), given the level/ranges of organizational pressure to close revenue gap, maximize profitability and minimize costs?. Please mark these ranges in the graph that best represents this relationship to your opinion”. Figure 6-11 shows the graphical form suggested to the experts.

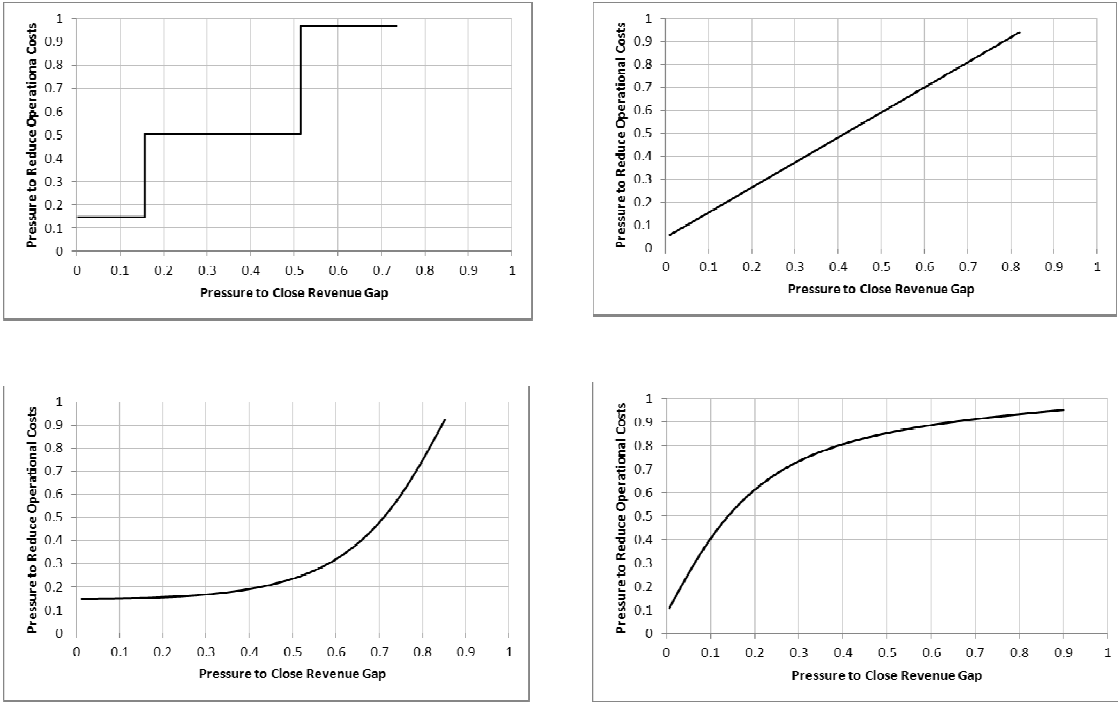


Figure 6-11. Pressure to close revenue gap/pressure to reduce operational costs relationship; format options

From aggregating experts' opinions, the following step function in Figure 6-12 was derived to represent the relationship between "Pressure to Reduce Revenue Gap" and "Pressure to Reduce Operational Costs".

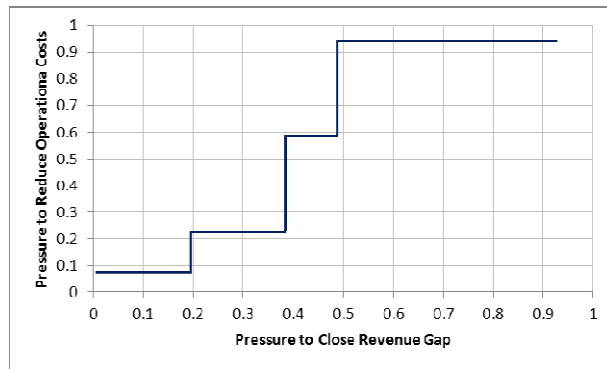


Figure 6-12. Expert opinion; Pressure to close revenue gap/pressure to reduce operational costs relationship

The closest parametric function to this step function is a negative exponential function in the form of;

$$PROC = (1 - \text{Exp}(-p_6 \times PCRG)) + p_7$$

Where,

PROC: Pressure to Reduce Operational Costs

PCRG: Pressure to Reduce Revenue Gap

and p_6, p_7 are model parameters.

Fitting the negative exponential function to experts' assessments reveals the optimum values for the parameters as;

$$p_6 = 1.79$$

$$p_7 = 0.077$$

Figure 6-13 shows the function for “Pressure to Reduce Operational Costs” using the optimum values for the parameters.

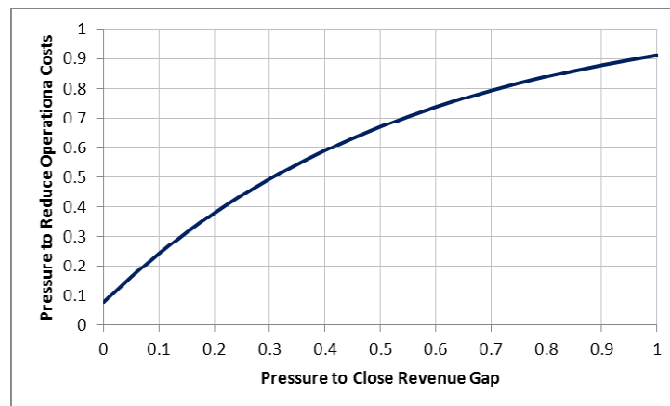


Figure 6-13. Pressure to close revenue gap/pressure to reduce operational costs relationship; functional form

4. Willingness/Ability to Invest in Proactive Safety Interventions

Given that willingness and ability to invest proactively in safety could technically be measured with the actual dollar value makes it a tangible concept in principal. For instance the records ideally show how much money is spent each year on safety interventions, and there is a direct relationship between willingness to invest proactively in safety and the actual dollar amount that get invested. So establishing the relationship between pressure to close revenue gap and willingness/ability to invest in safety should technically be straightforward. The challenge is though, that there are other reasons for investing in safety interventions; some are mandatory, and from data, it is very difficult to disentangle annually what investment is voluntary and what is being driven by some external regulatory body. Regulatory authorities (e.g. state's department of health), sometimes requires hospitals to demonstrate action on a certain, non-negotiable safety activities, additionally the certification body (i.e. The Joint Commission on Accreditation of HealthCare Organizations) then they would like hospitals to have a couple of optional investments and that is up to the hospitals to decide on what safety aspect they want to invest. These investments maybe in response to a finding by a third body payer (e.g. CMS) as well. It is very difficult to find out what part of the investment has been elective.

Due to these challenges and also difficulties to extract quality accounting data that differentiates the types of safety investments, and additionally given that only 9 years of financial data was available to us in this study, it is difficult to derive the relationship between pressure to reduce revenue gap, and ability to proactively invest in safety empirically. Table 6-2 shows the estimated safety investments and the operating margin for years 2003-2011(projected).

Year	OM	OM Dollars	Estimated Safety Investment
2003	-0.032	-\$14,600,000.00	\$150,000.00
2004	0.012	\$13,600,000.00	\$100,000.00
2005	0.034	\$33,700,000.00	\$350,000.00
2006	0.026	\$27,500,000.00	\$200,000.00
2007	0.047	\$52,388,770.15	\$450,000.00
2008	-0.009	-\$10,253,429.78	\$400,000.00
2009	0.013	\$14,847,535.95	\$200,000.00
2010	0.021	\$24,044,442.35	\$200,000.00
2011	0.035	\$40,174,255.76	\$200,000.00

Table 6-2. Operating margin and safety investment data, 2003-2011

In fact the regression analysis shows a very weak correlation between the operating margin and the actual dollars invested in safety (Figure 6-14), which we argue that is due to the lack of quality data, as was discussed.

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.65565673							
R Square	0.429885748							
Adjusted R Square	0.334866706							
Standard Error	99288.59757							
Observations	8							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	44600646353	4.46E+10	4.524206	0.077528042			
Residual	6	59149353647	9.86E+09					
Total	7	1.0375E+11						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	217578.4857	40967.18893	5.311043	0.001811	117335.3856	317821.5858	117335.3856	317821.5858
X Variable 1	3208679.594	1508534.172	2.127018	0.077528	-482570.55	6899929.737	-482570.5496	6899929.737

Figure 6-14. Operating margin and safety investment; regression analysis

To elicit this relationship from our experts, given that the relationship between financial situation of the hospital and investments made proactively in safety is a tangible relationship, we asked our experts:

“Given your assessment on “Pressure to Close Revenue Gap”, as a function of “Operating Margin”, what is the likelihood of observing change (increase or decrease) in the level hospital’s willingness and ability to invest proactively in safety, given the level/ranges of organizational pressure to close revenue gap, maximize profitability and minimize costs?. Please mark these ranges in the graph that best represents this relationship to your opinion”.

Figure 6-15 shows the graphical form suggested to the experts.

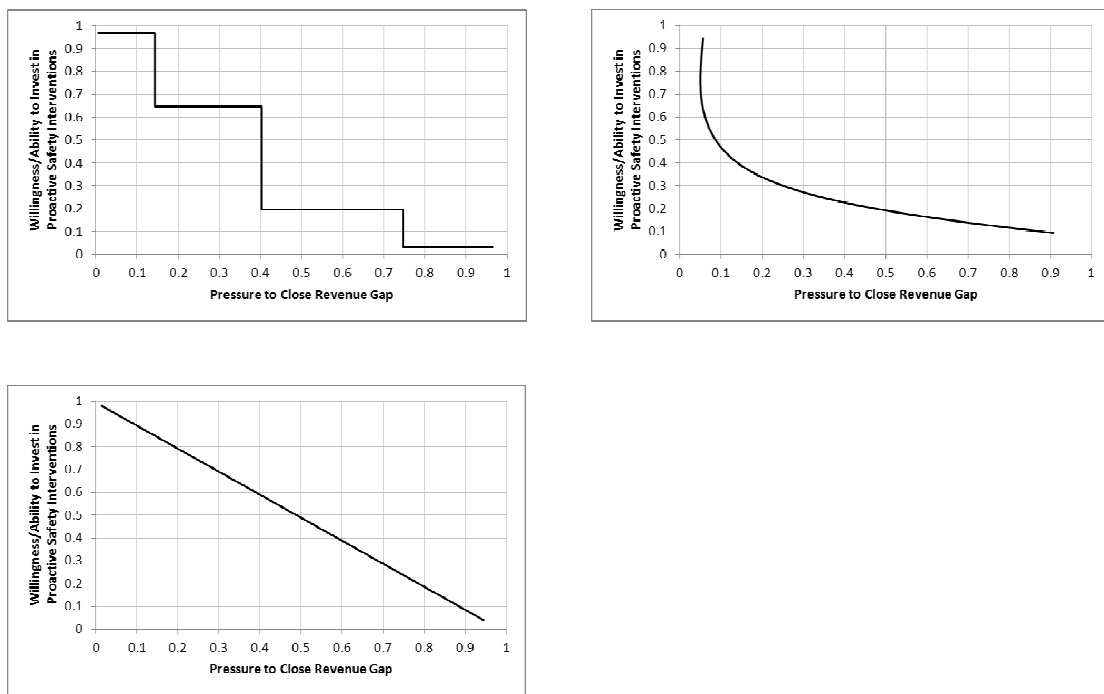


Figure 6-15. Pressure to close revenue gap/willingness ability to invest in safety; format options

Experts, almost unanimously picked the decreasing exponential format, but with a twist that for low pressures on closing revenue gap, the willingness to invest in safety is still very high because of the potential cost saving effects and partly the regulatory supervisions and then starts a dramatic decline, hence we have modeled this node as:

$$WIPS = \text{Exp}(-p_9 \times (PCRG - p_8))$$

Where;

WIPS: Willingness/Ability to Invest in Proactive Safety Interventions

PCRG: Pressure to Close Revenue Gap

and p_8, p_9 are model parameters.

Fitting the point estimates provided by experts on the graph, reveals the values of parameters:

$$p_8 = 0.2$$

$$p_9 = 22.687$$

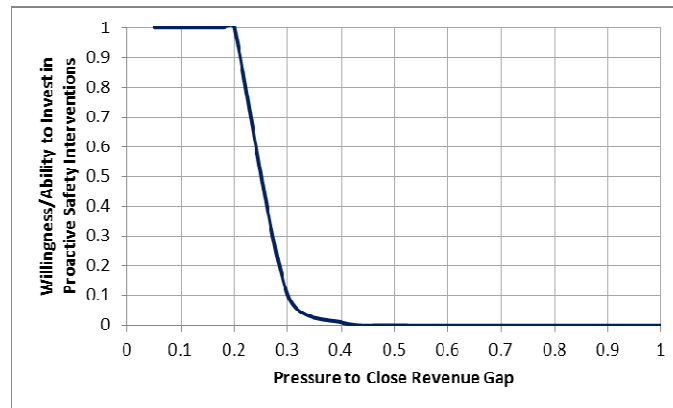


Figure 6-16. Expert opinion; pressure to close revenue gap/willingness ability to invest in safety relationship

6.2.2 Data Driven Factors

In this section, we will discuss the formulation and quantification of other factors in the model, including model constants that have been derived from clinical data.

5. *Probability of LOS Too Short to Meet All Needs*

Under high pressure to reduce LOS, not all patients are discharged prior to complete readiness, but the likelihood of a patient being discharged before all his/her needs are met increases. To capture this probabilistic notion in the model, we can use “*Readmissions*” within 24, 48 or 72 hours of discharge as a proxy. Readmission within 72 hours is probably a safe assumption that the patient was discharged inappropriately. That being said, there is still much more to it due to patient’s conditions complexity. For instance patients with Congestive Heart Failure, CHF, (where heart can’t pump enough blood to the organs) and the difference between their hearts maintaining the steady state versus not functioning, is a very fine line. Those patients may be admitted and stay in the hospital for two days and receive care, and they will be back to the hospital in 4-5 days, and this downward spiral continues till they are deceased. These types of discharges could not be categorized as inappropriate, because their stay in hospital, after they receive care for the first admission, is no longer justifiable, but you know that they will be back within a few days.

We extracted data for 70419 admissions, for years 2006 and 2008. There is a national average for LOS, for each diagnosis related group (DRG code). That means, every patient is assigned a DRG code (for his/her major diagnosis) upon admission. Each patient record we acquired indicates his/her diagnosis code, the national average LOS, the actual

LOS of that patient, and of course patient identifier. We also extracted data on all admission and discharge dates and times (sometimes multiple admissions and multiple discharges).

Comparing the national average LOS for each patient case, and the actual LOS of that patient reveals whether or not a patient was discharged earlier than anticipated, and by how many days (hours). At this stage we also eliminated the patient records where the patient had been deceased. Next, for the patients that shoed an early discharged we checked whether they had been readmitted to the hospital in the next 24, 48 and 72 hours of the previous discharge. This enables us to calculate the likelihood of a patient being readmitted to the hospital, due to an earlier than anticipated discharge and could be used as an indicator that the patient’s LOS has been too short to meet all his or her needs. Table 6-3 shows the average probability of LOS being too short to meet all patient needs, based on readmissions within 24, 48 and 72 hours.

	Number of Patients Readmitted	Probability of LOS Too Short
Readmission within 72 Hrs	615	0.091300475
Readmission within 48 Hrs	415	0.061609264
Readmission within 24 Hrs	227	0.033699525
Total Early Discharge	6736	

Table 6-3. Probability of LOS too short to meet all patient’s needs

Additionally, assuming readmission within 72 hours as the indicator of an early and inappropriate discharge we can obtain the distribution of the probability of readmission within 72 hours (3 days), as a function of how early the patient has been discharged compared to the national average of LOS for patient’s diagnosis code. Figure 6-17 represents this distribution. The data represents a beta distribution with parameters,

$a = -6.767E=6$ and $b = -0.0026$.

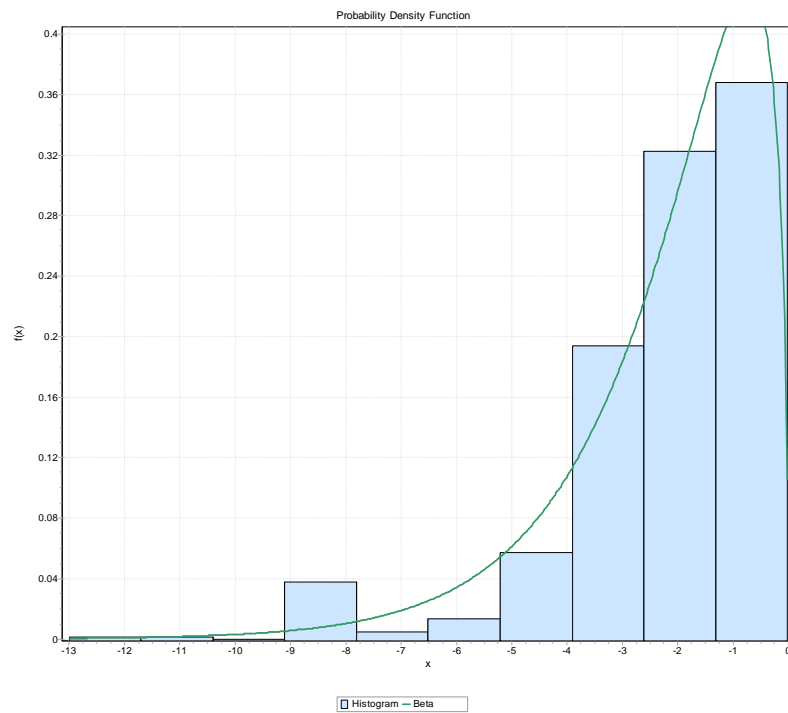


Figure 6-17. Number of days in inappropriate early discharge; pdf

Figure 6-18 depicts the CDF for this distribution.

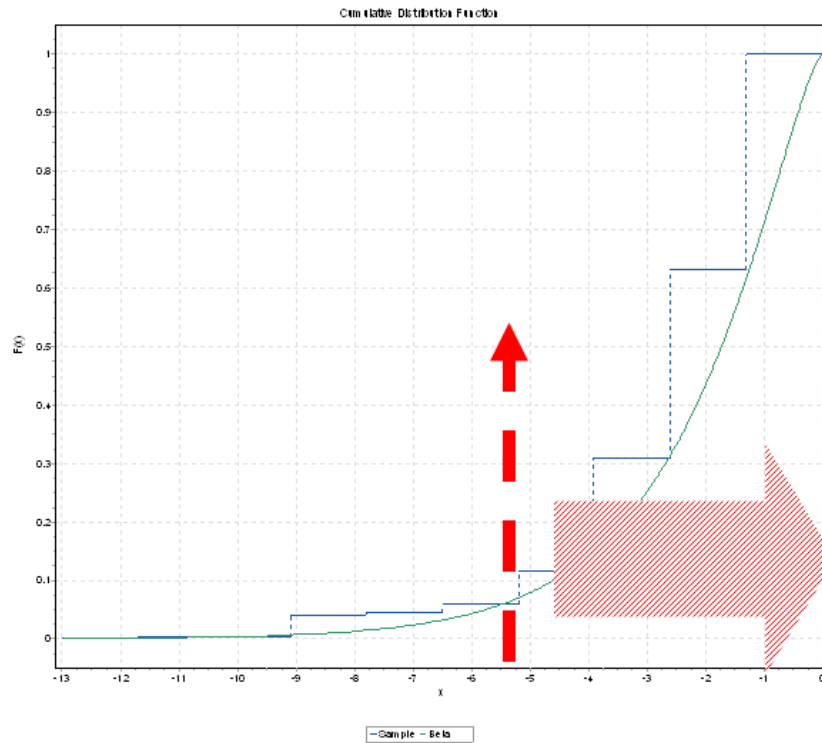


Figure 6-18. Number of days in inappropriate early discharge; cdf

For instance, if a patient is discharged 5 days or less, earlier than the national average of LOS for the specific DRG code, there is over 90% chance (the area under the curve, shown with the dashed arrow in Figure 6-18, that he or she will be readmitted to the hospital within 72 hours.

6. Probability of LOS Being Longer Than Needed

Some types of adverse events may increase in likelihood of occurrence, due to prolonged LOS, simply because the patient's exposure to the risk increases. For instance prolonged LOS may increase patient's chances of experiencing pressure ulcer because he/she will spend more time in bed. Another example where prolonged

LOS increases the risk of adverse event, are falls. The more the patient stays in the hospital, the higher the chances of him/her falling.

Although, the lack of pressure to optimize LOS, at least commonly, cause patients to stay in the hospital longer than they really need, but experts picture situations where patients do over stay in the hospital.

For instance, if the hospital admits a homeless patient and physicians start insulin treatment because the patient has been diagnosed with diabetes, they cannot be discharged to the street. Sometimes those patients have to stay till hospital figures out where they are going to send the patient.

Let's look at another scenario for prolonged LOS. After patient has received care for the major diagnosis he or she has been admitted for, if there is still need for more care that is not specialty care, the patient will be sent to an "*Extended Care Facility*", such as a nursing home. If nursing homes are totally full and have no room the patient has to be waiting while kept in the hospital for an opening in the nursing home. More often than not, nursing homes refuse to accept a patient due to his/her certain condition (e.g. MRSA: methicillin Resistant Staphylococcus Aureus, a bacteria resistant to antibiotics) to avoid its spread to everyone else in the nursing home, and hospital has to find a nursing home that is willing to accept the patient, while the patient stays in the hospital.

Other than these situations, if patients stay for a long periods of time in the hospital (and hence, increase their exposure to the risk of certain adverse events), it's because

based on their condition, that is how long it takes for them to receive the care they need, and this long length of stay is not affected by low pressure to optimize LOS.

This is very difficult to model and no data is available that directly indicates what are the chances that people will overstay because at the hospital level, there is no pressure to reduce LOS. To include this in the model though, we had to come up with a proxy.

Records show that it usually takes longer to discharge patients on weekends than in week days (where hospital is working at 96-98% capacity, and there is a lot of pressure to discharge as many patients as possible, if it is appropriate). Our experts contributed this to the fact that access to “services” (e.g. oxygen) is a little more difficult in weekends than in weekdays, but not much and low pressure to optimize LOS will result in patients staying over the weekend due to these small challenges in arranging these services. The pressure to reduce LOS is greatest in weekdays, because people come in for elective surgeries and procedures the next day and the hospitals really needs the beds, and this pressure to reduce LOS relaxes a bit during the weekend.

We extracted 77403 patient records from 2003 to 2005, and compared their LOS to the national average of LOS for the primary diagnosis code (DRG code). Table 6-4, Shows a few records of this data. 30665 patients out of this population had LOS greater than that of the national average. Hence we can estimate the probability of prolonged LOS (for any reason) to be, $\frac{30655}{77403} = 0.39$. On the other hand we know that not of this prolonged LOS is due to lower organizational pressures to discharge patients, but rather due to patient’s medical conditions.

DRG Code	Actual LOS	National Average LOS	Difference in Actual and Average LOS	Prolonged LOS (Yes=1, No=0)	LongLOS_Condition: Pressure Ulcer	LongLOS_Condition: Fall	LongLOS_Condition: Infection	LongLOS_Condition: Medication Error
35	48	60	-12	0	0	0	0	0
35	48	60	-12	0	0	0	0	0
35	48	60	-12	0	0	0	0	0
35	144	60	84	1	1	1	1	1
35	24	60	-36	0	0	0	0	0
35	48	60	-12	0	0	0	0	0
35	24	60	-36	0	0	0	0	0
37	72	62.4	9.6	1	0	0	0	0
37	48	62.4	-14.4	0	0	0	0	0
37	48	62.4	-14.4	0	0	0	0	0
37	360	62.4	297.6	1	1	1	1	1
37	48	62.4	-14.4	0	0	0	0	0
37	24	62.4	-38.4	0	0	0	0	0
37	24	62.4	-38.4	0	0	0	0	0
40	24	60	-36	0	0	0	0	0
40	24	60	-36	0	0	0	0	0
40	144	60	84	1	1	1	1	1
44	48	98.4	-50.4	0	0	0	0	0
44	48	98.4	-50.4	0	0	0	0	0
44	24	98.4	-74.4	0	0	0	0	0
44	96	98.4	-2.4	0	0	0	0	0
44	48	98.4	-50.4	0	0	0	0	0
44	72	98.4	-26.4	0	0	0	0	0
44	48	98.4	-50.4	0	0	0	0	0
44	72	98.4	-26.4	0	0	0	0	0
44	192	98.4	93.6	1	1	1	1	1

Table 6-4. Sample data records extracted to indicate whether the patient had prolonged LOS

Dierks (unpublished data, 2011), has conducted an study of the patients with prolonged LOS (39% of hospital population), which indicates that all other patients' conditions being equal (relatively), 11%-17 % of these patients with prolonged LOS, have been those with weekend/Monday AM discharges. This makes for 4.3 %to 6.6%, of total hospital population, having prolonged LOS because the time of discharge is a weekend where the pressure to optimize LOS is lowest. We will use the middle of this range for our model quantification (5.45% probability of prolonged LOS due to low level of pressure to optimize LOS). In performing uncertainty analysis, we will assume a normal distribution: Normal($\mu = 0.045, \sigma = 0.38$) , for this value.

6.2.3 Magnitude of the Effects of Cost Cutting Strategies on the Risk of Adverse Events

As discussed in section 6.1, the strategies that decision makers take in response to an unfavorable revenue, will affect the risk of certain adverse events. In this section we discuss the magnitude of these effects.

7. Magnitude of Change in the Risk of Adverse Event Due to Shortened LOS

To formulate this node in the model, we have used a combination of subjective and data driven approach. We argued that at certain high levels of pressure to reduce LOS, there is a chance that some patients will be discharged prior to readiness, and we used readmission within 72 hours data, to calculate the probability of early discharge. To find out what levels of pressure may cause an early discharge; we used our experts' subjective opinion, using interview guides in appendix D. To elicit this information from the experts the following questions were asked:

“As was discussed in other rounds of interview, the pressure to optimize LOS, may affect risk of adverse events in two ways, first, it may increase the probability of experiencing an adverse events, because some patient’s LOS may be too short to meet all his/her needs. Second, it may reduce the probability of some adverse events because it simply reduces the exposure and if the pressure to optimize LOS is too small, some patients may stay in the hospital longer than they really need and be exposed to certain adverse events.

- *To your experience at what level (range) of pressure to optimize LOS, we might start to see the effect of this pressure on risk, because the pressure is too high that some patients may be discharged a bit prematurely?”*

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1

Table 6-5 shows experts’ responses and assessments.

Expert	Pressure to Optimize LOS
Expert 1	0.8
Expert 2	0.4
Expert 3	0.3-0.5
Expert 4	0.8
Expert 5	0.3
Expert 6	0.4-0.5
Expert 7	0.8-0.9
Expert 8	0.8-0.9
Expert 9	0.7-0.8
Expert 10	0.8-0.9
Expert 11	0.6-0.7
Expert 12	0.4-0.5
Expert 13	0.4-0.5
Expert 14	0.6-0.7
Expert 15	0.6-0.7
Expert 16	0.8

Table 6-5. Expert opinion; level of pressure to optimize LOS triggering early discharge

To aggregate experts’ inputs, we will use the Bayesian method discussed in 4.3.3.2.1, which results in a posterior distribution shown in Figure 6-19 with mean 0.6 and variance 0.01.

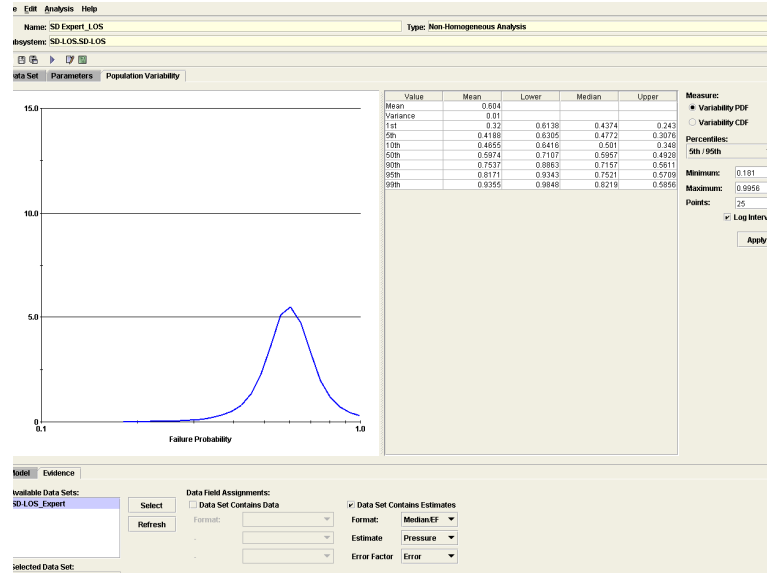


Figure 6-19. Experts' opinion on level of pressure to optimize LOS triggering early discharge; aggregated

We will use the mean value for the quantification of the model, and the variability for uncertainty analysis. Now given that the pressure to optimize LOS is at a level that may cause some patients to be discharged early, we will have to determine by what magnitude does their risk, of specific adverse event changes. Dierks et al. (manuscript in preparation, 2011), have conducted a study that measures the magnitude of change in the risk of certain adverse events, not reimbursed by third party insurers, due to: Shortened LOS, Prolonged LOS, and Understaffing.

The estimates provided by this study, shown in Table 6-6 have been used in our model quantification. According to this study, shortened LOS decreases the baseline risk of pressure ulcer and line infection by 3% and 5% respectively, due to reducing exposure.

Table 1: Hospital-Acquired Conditions for FY 200x							
Selected HAC	CC/MCC (ICD-9-CM Codes)	Number of cases	Denominator or	Exposure Risk Requirement(s)	LOS- or Time-Dependent Adverse Event?	Resource-Dependent Adverse Event?	Production-Pressure-Dependent Adverse Event?
					Magnitude of Change in Risk with Shortened LOS	Magnitude of Change in Risk with Staffing?	Magnitude of Change in Risk with Increased Production Pressure?
Foreign Object Retained After Surgery	988.4 (CC); 988.7 (CC)	15	24843	Must have had surgical procedure	0	0	0.15
Air Embolism	989.1 (MCC)	0	11989	Must have a large bore central venous catheter or venous interventional procedure	0	0	0.15
Blood Incompatibility	988.0 (CC)	0	4327	Must receive blood product; function of number of different blood products received	0	0	0
Pressure Ulcer Stages II and IV	707.23 (MCC); 707.24 (MCC)	6	10000	Advanced age and immobility	-0.03	0.25	0
Falls and Trauma: Fracture, Dislocation, Intracranial Injury, Crushing Injury, Bum, Elastic Straps	Codes within these ranges on the CC/MCC Lists: 800-829, 830-838, 850-854, 925-929; 930-949; 991-994	23	49000		See comments	See comments	0
Catheter-Associated Urinary Tract Infection (UTI)	565.51 (CC); 565.52 (MCC) (ICD-9-CM codes for UTI from being used in ICD-10) 560.10 (CC); 560.11 (MCC); 560.2 (MCC); 560.3 (CC); 560.31 (MCC); 560.32 (MCC); 560.33 (MCC); 560.34 (MCC); 560.35 (MCC); 560.36 (MCC); 560.37 (MCC); 560.38 (MCC); 560.39 (MCC)	125	18754	Must have urinary catheter in place; function of duration of catheter in place (some other factors, as well)	-0.05	0.05	0
Vascular Catheter-Associated Infection	999.31 (CC)	67	11983	Must have vascular catheter in place; function of duration of catheter in place (some other factors, as well)	-0.05	0.05	0
Manifestations of Poor Glycemic Control	249.10-249.11 (MCC); 249.20-249.21 (MCC); 250.10-250.13 (MCC); 250.20-250.23 (MCC); 251.0 (CC)	374	12919	Must have underlying insulin resistance; function of dietary orders, as well	-0.05	0.05	0
Surgical Site Infection, Mediastinitis, Following Coronary Artery Bypass Graft (CABG)	519.2 (MCC); And one of the following procedure codes: 39.10-39.13	2	815	Must have CABG surgical procedure	-0.05	0.05	0.05
Surgical Site Infection Following Certain Orthopedic Procedures	868.87 (CC); 868.89 (CC); And one of the following procedure codes: 31.01-31.08, 31.29-31.24, 31.31-31.38, 31.83, or 31.85	11	955	Must have certain orthopedic procedures surgical procedure	-0.05	0.05	0.05
Surgical Site Infection Following Bariatric Surgery for Obesity	868.87 (CC); 868.89 (CC); And one of the following procedure codes: 41.14-41.16, 41.23-41.24, 41.31-41.36, 41.37-41.38	11	955	Must have bariatric surgical procedure	-0.05	0.05	0.05
Deep Vein Thrombosis and Pulmonary Embolism Following Certain Orthopedic Procedures	415.11 (MCC); 415.19 (MCC); 423.40-423.42 (MCC); And one of the following procedure codes: 30.05-30.07, 31.31-31.32, 31.34	2	432	Must have certain orthopedic procedures surgical procedure	-0.02	0.15	0

Table 6-6. Magnitude of change in the risk of hospital acquired adverse events due to staffing, prolonged and shortened LOS

8. *Magnitude of Change in the Risk of Adverse Event Due to Prolonged LOS*

Similar to part 7, we elicited experts’ opinion on the levels of pressure to optimize LOS that corresponds to changes in risk of adverse events due to prolonged LOS.

“To your experience at what level (range) of pressure to optimize LOS, we might start to see the effect of this pressure on risk, because the pressure is too low that some patients may stay longer than they really need to which may increase their risk of being exposed to and experiencing certain adverse events?”

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1

Table 6-7, reflects experts’ responses and assessments.

Expert	Pressure to Optimize LOS
Expert 1	0.2
Expert 2	0.3
Expert 3	0.1-0.2
Expert 4	0.2
Expert 5	0.1
Expert 6	0.2-0.3
Expert 7	0.1
Expert 8	0.1
Expert 9	0-0.1
Expert 10	0.1-0.2
Expert 11	0.1-0.2
Expert 12	0-0.1
Expert 13	0-0.1
Expert 14	0-0.1
Expert 15	0
Expert 16	0

Table 6-7. Expert opinion; level of pressure to optimize LOS triggering prolonged LOS

Aggregated experts’ assessments, shown in the distribution below (Figure 6-20), has a mean of 0.11 and variance 0.03.

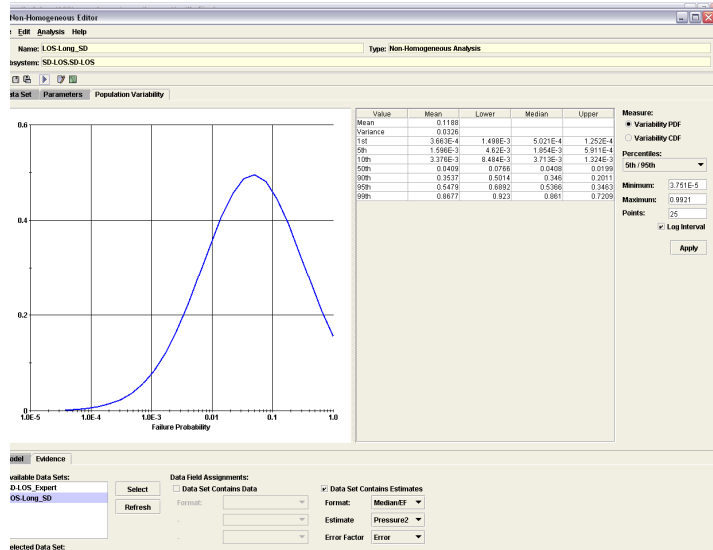


Figure 6-20. Experts' opinion on level of pressure to optimize LOS triggering prolonged LOS; aggregated

The estimates provided by Dierks et al. (manuscript in preparation, 2011), shown in Table 6.6, have been used in our model quantification. According to this study, prolonged LOS increases the baseline risk of pressure ulcer and line infection by 25% and 5% respectively, to reducing exposure.

9. Magnitude of Change in Risk of Adverse Event Due to Understaffing

Dierks et al. (manuscript in preparation, 2011), estimate the magnitude of change in the risk of certain adverse events due to understaffing, shown in Table 6.6. But the

characteristics of “being understaffed” or “staff adequacy” are not clearly defined in the literature, and have not been modeled and formulated before. The following is a discussion on our approach to characterize and model “understaffing” and “staff adequacy”, probabilistically using patient complexity scores.

- *Complexity Score Distribution and Probability of Understaffing*

Determining the probability of a unit being understaffed is not a trivial task. The notion of nurse to patient ratio, which is a mandatory ratio to be maintained by hospitals, is usually maintained by hospitals due to consequences imposed by regulatory authorities. Experts believe that the concept of understaffing and staff adequacy goes way beyond the nurse-patient ratio. The staffing ratio only tells you how many nurses with a RN degree you need to have given the census on the hospital floor, but no two patients are exactly equivalent and sometimes there is a sudden increase in intensity of the workload where although you might be having the required staffing ratio but the adequacy of staffing drops because of that. Since there is not clinical data or reliable data for that matter, exists of adequacy of staffing we had to find a measure that best represents the concept of staff adequacy. For this purpose, we have utilized “*Case Mix Index (CMI)*”, which is the average Diagnosis Related Group (DRG) weight for all of a hospitals Medicare patients, and can be used as an indicator of patient’s complexity of illness (Steinwald and Dummit, 1989).

The basic idea is, that a combination of high pressure to reduce operational costs (i.e. staffing) and high level of patient complexity, will lead to inadequate staffing situation.

We have extracted patient complexity score data, for 970 days (from August 31st 2008, to April 29th 2011), for 33 units in the hospital. From this data we have calculated the mean and the median of complexity scores for each floor, each day. What is challenging though using the mean is that, the average value of complexity can be calculated for each floor but the complexity score is not normally distributed for each floor. For instance if you have 10 patients on 1 floor, on any given day, we may have 9 patients with complexity of 1, and one patient with complexity of 18. The average of these complexities for this floor at this day is 2.7. On the other hand a floor with 10 patients that all have complexity of 2.7, also gives an average complexity of 2.7. What is unclear is whether complexity composed of 9 easy patients and one difficult patient is different from average complexity of 2.7, where everyone has a complexity index of 2.7.

However, to characterize the workload based on the complexity of the patients more realistically, we will use the median complexity score for each floor, each day. Assuming homogeneity of data from all 34 units, the complexity score distribution for all units (34 units), and all days (970 days) best fits a Weibull distribution, with parameters $\alpha = 4.87$, $\beta = 1.832$, (mean= 1.745, and variance= 0.287) which is used as a representative of the distribution of complexity scores hospital wide. Figure 6-21, shows this distribution, and Table 6-8, contains the descriptive statistics of this distribution.

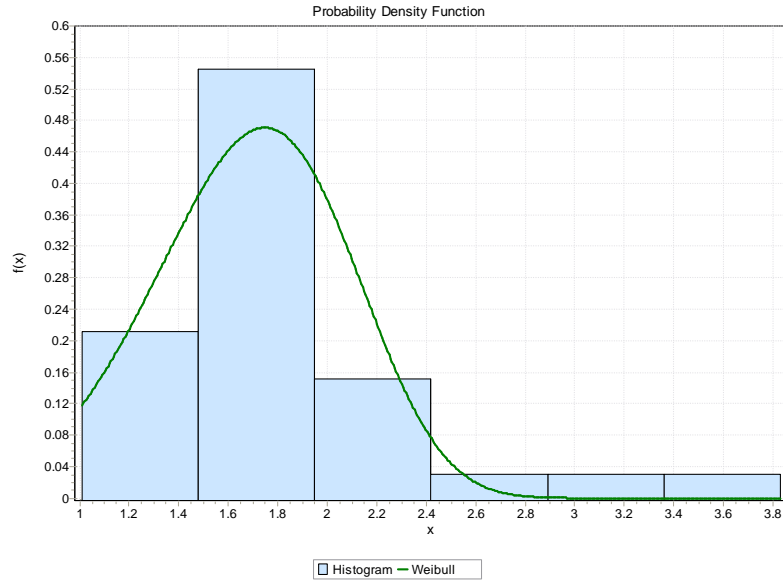
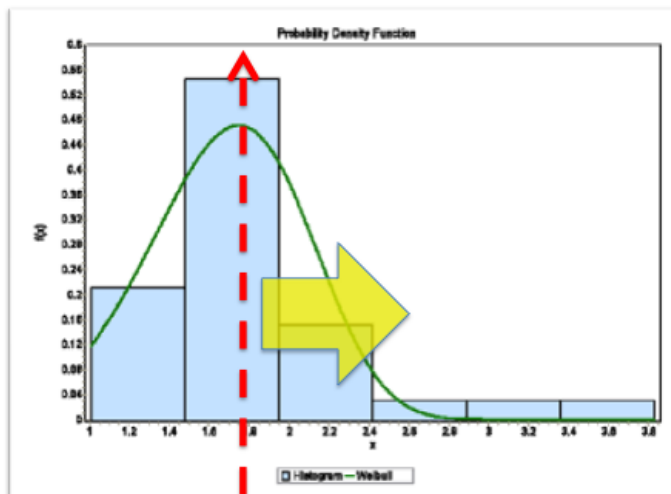


Figure 6-21. Case-mix complexity distribution across hospital

Statistic	Value	Percentile	Value
Sample Size	33	Min	1.0089
Range	2.8224	5%	1.0321
Mean	1.7448	10%	1.2711
Variance	0.28686	25% (Q1)	1.4834
Std. Deviation	0.53559	50% (Median)	1.5992
Coef. of Variation	0.30696	75% (Q3)	1.9055
Std. Error	0.09323	90%	2.3679
Skewness	2.2682	95%	3.2381
Excess Kurtosis	6.987	Max	3.8313

Table 6-8. Case-mix complexity distribution across hospital; descriptive statistics

In presence high levels of pressure to reduce operational costs (due to unfavorable revenue gap), a unit with high patient complexity. Using experts' opinion, we can formulate a combination of level of patient complexity and pressure to reduce operational costs that may lead to inadequacy in staffing, probabilistically.



C

Figure 6-22. Calculating the probability of understaffing

For instance, when pressure to reduce operational costs is “P” (given by experts), and complexity score of the unit is at “C”, the probability of the unit being understaffed is the area under the curve in

Figure 6-22, to the right of the complexity level “C”.

To determine combinations of complexity scores and cost reducing pressures that may cause understaffing, experts were asked to provide their opinion through the following interview question;

“- The pressure to reduce operational costs, may affect risk of adverse events due to understaffing. Assume that when there are pressures to cut operational costs, the organization may respond by reducing numbers of staff or staffing with less experienced staff (at a lower cost).”

- *To your experience what level (range) of pressure to reduce operational costs are great enough such that some patients may experience an adverse*

event because the unit is not sufficiently staffed (either due to lower numbers of staff, or lower quality of staff?)

*You may use an arrow to indicate a precise point or circle one of the ranges above to indicate a broader range estimate.

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1

The idea is that the impact of lower staffing numbers and/or less experienced staff may depend on the complexity of the case mix. With this assumption in mind, suppose we take the average of the complexity scores across all units in the hospital and all inpatient days. Please indicate in the table below, where the pressure to reduce operational costs/staffing begins to influence the probability of an adverse event as a function of the complexity of the patient population.

Highest Complexity	High Complexity Score: Greater than 2.2																												
	E5, E10	E2, E3, E4, E6 E7, E9, E14	E11		E1																								
	E5, E13	E2, E3, E4, E6, E7 E8, E10, E13, E14	E11, E13		E1																								
	E5	E2, E3, E4, E6, E7 E14	E11		E1																								
Lowest Complexity	Medium Complexity Score: Between 1 and 2.2																												
	E5		E2, E4, E6	E3, E14	E1, E3, E7, E8 E10, E11	E3	E3																						
			E2, E12	E4, E5, E6, E14 E13	E11, E13	E1, E8, E10, E13	E7	E3																					
			E2	E4, E5, E6, E12 E14	E11	E1	E8, E10	E3, E7																					
				E2, E4, E5	E6	E1	E8	E7, E10 E14	E3																				
				E5	E2	E5, E6	E1, E4, E13	E7, E8, E14, E11 E13	E10, #13	E3, E13																			
						E2, E5	E6	E1, E4, E7, E8 E9, E12, E14, E11		E3, E10																			
Low Complexity Score: Less than 1																													
Lowest Complexity																													
<table border="1" style="width:100%; text-align:center;"> <tr> <td>0-0.1</td> <td>0.1-0.2</td> <td>0.2-0.3</td> <td>0.3-0.4</td> <td>0.4-0.5</td> <td>0.5-0.6</td> <td>0.6-0.7</td> <td>0.7-0.8</td> <td>0.8-0.9</td> <td>0.9-1</td> </tr> <tr> <td colspan="5">Lowest Pressure = 0</td> <td colspan="5">Highest Pressure = 1</td> </tr> </table>										0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Lowest Pressure = 0					Highest Pressure = 1				
0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1																				
Lowest Pressure = 0					Highest Pressure = 1																								

Table 6-9. Experts' opinion; combination of level of pressure to reduce operational cost, and complexity score effecting the probability of adverse events due to understaffing

Each expert E_i , marked in a table like Table 6-9, what combinations of organizational pressure to reduce operational costs and staff cuts, and complexity of patient population may start to influence the probability of a patient experiencing an adverse event, due to understaffing. Table above reflects the collective experts' assessments.

To use these assessments in the model, we have discretized the space shown in Table 6-9 above, into three bounds, in a way that covers as many expert assessments as possible:

- A. Low Pressure- High Complexity: where pressure to reduce operational cost is between 0 and 0.3, and complexity is above 2.2 (yellow area).
- B. Medium Pressure- Medium Complexity: where pressure to reduce operational cost is between 0.3 and 0.6, and complexity is between 1 and 2.2 (orange area).
- C. High Pressure- Low Complexity: where pressure to reduce operational cost is greater than 0.6 and complexity is less than 1 (red area).

If the combination of the level of pressure to reduce operational costs and complexity of patient population falls in any of the above; A, B or C categories, we assume that there is a certain probability (determined in Table 6.6) that the patient will experience a certain adverse event (in the case of our study, pressure ulcer or line infection), due to understaffing (of various forms such as inadequate number of staff, inadequate experience of staff, etc.).

With the three critical limits for low, medium and high complexity, we can obtain the probability that complexity exceeds these limits from the CDF of case-mix complexity distribution in Figure 6-21. This CDF can be seen in Figure 6-23.

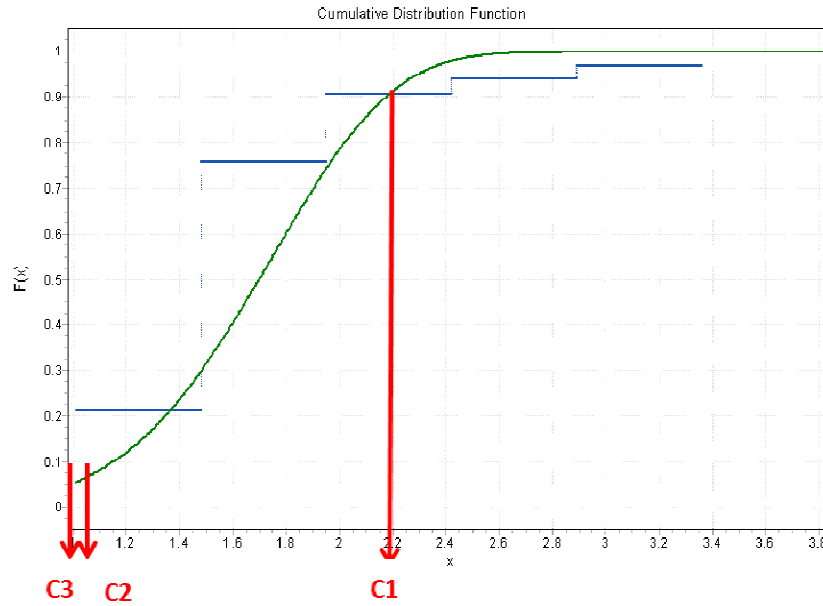


Figure 6-23. Case-mix complexity distribution across hospital; CDF

Hence:

Probability of complexity exceeding $C1=2$ is 0.1

Probability of complexity exceeding $C2=1$ is 0.95

Probability of complexity exceeding $C3=0$ is 1

10. Magnitude of Change in Risk Due to the Lack of Investment in Safety Interventions

It is almost possible to distinguish how much of safety related investments directly affect which adverse event, and what the exact magnitude of change in risk of adverse events is contributable to how much of these investments. The data available in this study, on overall safety related investments and the relative frequency of the two adverse events we have studied in this research (pressure ulcer and line infection) does not reveal meaningful correlation between the two. Table 6-10 shows the prevalence of pressure ulcer and the expenditure on safety from 2003 to 2011, no meaningful information could be extracted regarding the effect of change in the investment and change in the prevalence of pressure ulcer.

Fiscal Year	Estimated Investment Safety	Prevalence of Pressure Ulcer
2003	\$150,000	0.030997057
2004	\$100,000	0.034305162
2005	\$350,000	0.039968817
2006	\$200,000	0.040111022
2007	\$450,000	0.04034664
2008	\$400,000	0.045892935
2009	\$200,000	0.045011043
2010	\$200,000	0.039947894
2011	\$200,000	0.031630898

Table 6-10.Safety investment data and prevalence of pressure ulcer

For example between years 2006 and 2007, there is 125% increase in safety investments, yet data shows 15% increase in the prevalence of pressure ulcer.

Experts’ opinion on magnitude of effect of ability to invest in safety on the risk of hospital-acquired conditions has been solicited through the following question and the aggregated results are reflected in Table 6-11.

“Assume that the pressure to reduce the revenue gap, will affect the level of willingness/ability to invest in proactive safety investments. The worse the financial situation gets, the less investments are made in safety programs. Assume that the more we spend on safety the less the chances of experiencing adverse events will be. If this willingness to invest in safety is a scale between 0-1 (0 meaning no ability/willingness to invest in safety, and 1 meaning highest level of ability/willingness to invest in proactive safety interventions);

- (a)Based on your experience at what level (range) of this willingness do we start to see changes in the risk of adverse event?

*You may use an arrow to indicate a precise point or circle one of the ranges above to indicate a broader range estimate.

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1
---------------	-------	---------	---------	---------	---------	---------	---------	---------	---------	-------	-----------------

Expert’s responses to this question are reflected in Table 6-11 below.

Expert	Ability/Willingness to Invest in Safety
Expert 1	0-0.1
Expert 2	0.2
Expert 3	0.2-0.3
Expert 4	0.1
Expert 5	0.2-0.3
Expert 6	0.3-0.4
Expert 7	0.7-0.8
Expert 8	0.2-0.3
Expert 9	0.5-0.6
Expert 10	0.3-0.4
Expert 11	0.1-0.2
Expert 12	0.2-0.3
Expert 13	0.1
Expert 14	0.1

Table 6-11. Experts' opinion; level of ability/willingness to invest in safety and the effect of risk of adverse events

Aggregating these estimates, using the Bayesian methods for expert assessments aggregation yields to the distribution below (Figure 6-24), with mean 0.22 and variance 0.15.

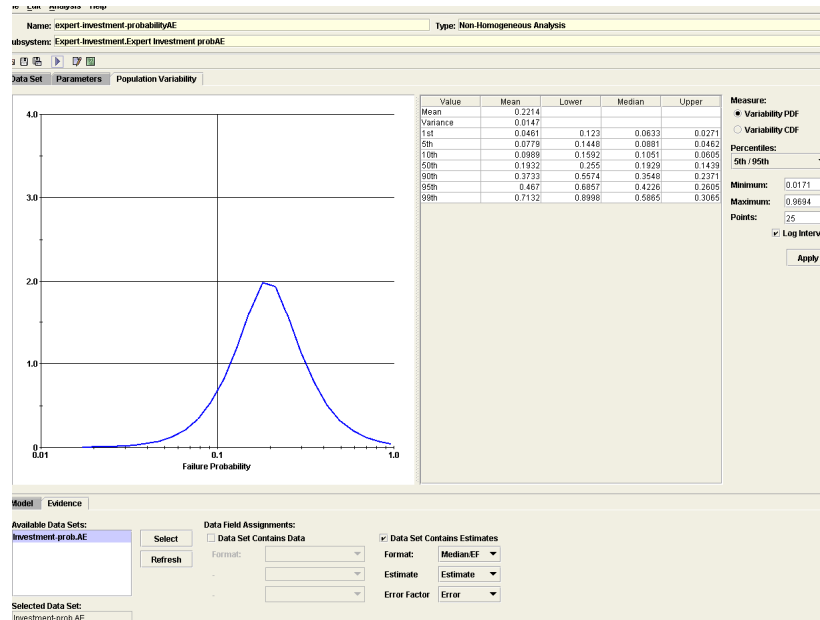


Figure 6-24. Experts' opinion on level of ability/willingness to invest in safety and the effect of risk of adverse events; aggregated

F. (b) Based on your experience, what is the magnitude of change in the risk of adverse events when there is an increase or decrease in investment in elective/proactive safety programs? Use the table below to indicate the relationship between changes in investments and magnitude of effect on risk of adverse events.

Experts' responses to this question, have been collected and summarized in Table 6-13 (E_i represents expert i). One of the experts did not provide an answer to this particular

question. Also, interestingly enough, three of our experts did not feel like the maximum limits of decrease and increase to the probability of adverse events suggested by us at the interview forms (Maximum 5% increase and decrease in the probability of adverse event), did justice to the magnitude of effect of safety investments on probability of adverse events. Hence they set up their own limits. For instance Expert 2 believed that high investments in proactive safety interventions can reduce the risk of adverse events by 25%, medium size investments could decrease the risk by 15%, and low investments could double the risk and increase it 100%. To aggregate experts' opinion, in assessing the magnitude of change in the risk of adverse events due to lack of investment in proactive safety interventions, we have discretized the space in Table 6.13, into 3 categories:

- A. High Ability/Willingness to Invest Proactively in Safety (yellow area)
- B. Medium Ability/Willingness to Invest Proactively in Safety (orange area)
- C. Low Ability/Willingness to Invest Proactively in Safety (red area)

In each of these areas, we aggregate experts' opinion by taking the weighted average, due to the lower number of experts (3 out of 13) who believed in much larger impacts of safety investments on risk of adverse events. The results of aggregation could be seen in Table 6-12 .

Willingness/Ability to Invest in Proactive Safety Interventions	Magnitude of Change in Risk of Adverse Event Due to lack of Safety Investments
High (>0.6)	-13%
Medium (>0.3 but <0.6)	-4%
Low (<0.3)	19%

Table 6-12. Experts' opinion on magnitude of change in the risk of adverse events due to lack of investment in safety; aggregated

6.2.4 Other Constants of the Model

In order to calculate the cost of caring adverse events, we need to find out how much impact a specific adverse event has on the LOS and how much longer patients will have to remain in the hospital to be treated for the adverse event they have experienced while in the hospital.

11- Increase in the LOS Due to the Adverse Event

1. Increase in the LOS Due to Pressure Ulcer

Using 627,595 patient records for years 2008-2010, we identified 86 admission diagnosis codes, common between pressure ulcer cohort (patients who did experience pressure ulcer while in the hospital) and non-pressure ulcer cohort (other patients who did not acquire pressure ulcer in hospital). Then, we averaged the LOS for all patients in the first (pressure ulcer cohort) and the second (non-pressure ulcer cohort) group, for each of the admission diagnosis codes. The difference between, average LOS of the two groups, naturally reveals how much longer do people who acquire pressure ulcer in hospital, will have to stay in the hospital compared to patients who did not experience pressure ulcer. Table 6-14, shows the average LOS for both cohorts, for each of the admission diagnosis group.

It is worthwhile noting that, for a couple of the admission diagnosis codes (which sometimes only had one patient), we realized that people who had experienced pressure ulcer had actually stayed shorter than people with the same diagnosis code in non-

pressure ulcer cohort. This is due to the fact that these patients are either discharged to other departments, or they are deceased.

Extracting this outlier data, the increase in the LOS due to experiencing pressure ulcer, is best presented by a Gamma distribution, with parameters $\alpha = 0.789$, $\beta = 20.134$, and mean 15.88 days.

Figure 6-25, shows this distribution. The mean of this distribution is days, which shows on average people who experience pressure ulcer stay 14 days longer in the hospital, which agrees with the studies we found in the literature. For instance Beckrich and Aronovitch (1999) found in a study that the increased in LOS due to pressure ulcer is between 14-17 days.

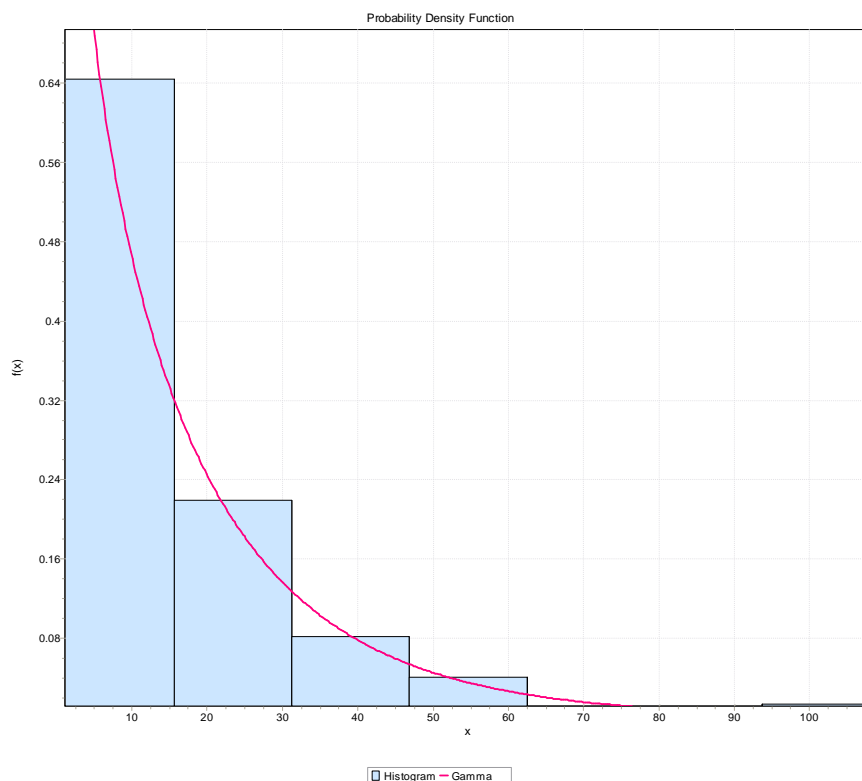


Figure 6-25. Distribution of increased LOS due to pressure ulcer

Pressure Ulcer Cohort		Non-Pressure Ulcer Cohort	
Admission Diagnosis Code	Average LOS	Admission Diagnosis Code	Average LOS
1508	18.00	1508	19.38
1510.00	35.00	1510	13.95
1518	24.00	1518	9.23
1550	123.00	1550	9.85
1570	10.00	1570	7.86
1830	13.00	1830	8.10
1840	4.00	1840	5.93
1915	40.00	1915	19.79
1983	10.33	1983	7.53
1991	15.00	1991	8.28
20023	43.00	20023	30.19
20050	21.00	20050	22.23
20280	47.00	20280	30.46
20300	131.00	20300	21.64
20500	76.67	20500	36.53
23876	106.00	23876	47.36
25013	12.00	25013	4.33
25060	35.00	25060	5.58
2761	7.00	2761	5.50
27651	9.00	27651	3.13
28800	12.00	28800	10.88
33524	17.00	33524	17.00
34982	23.00	34982	7.33
3968	43.00	3968	17.49
41041	41.67	41041	7.93
41071	20.40	41071	6.13
41401	49.00	41401	4.11
4240	16.00	4240	7.95
4271	7.00	4271	4.75
42741	14.00	42741	6.84
42843	7.00	42843	6.10
44023	18.00	44023	6.07
44024	17.50	44024	8.73
44101	36.00	44101	16.62
4416	10.00	4416	17.06
45341	35.00	45341	4.90
486	23.00	486	4.84
5070	16.50	5070	6.93
515	2.00	515	3.57
51881	29.00	51881	10.71
51902	20.00	51902	10.59
53561	3.00	53561	3.19
5551	38.00	5551	8.83
56081	17.00	56081	11.75
570	41.00	570	8.46
5712	51.00	5712	10.42
5722	37.00	5722	10.65
5724	27.00	5724	23.09
5770	61.00	5770	7.63
5793	28.00	5793	19.68
5849	18.00	5849	4.96
6822	3.00	6822	4.44
6826	12.00	6826	4.67
70719	7.00	70719	9.15
71107	18.00	71107	11.17
7213	8.00	7213	5.79
72271	23.00	72271	3.64
72888	6.00	72888	6.11
73313	12.00	73313	6.22
73315	8.00	73315	7.98
73730	14.00	73730	8.08
7907	27.00	7907	7.84
80121	27.00	80121	11.61
80126	21.00	80126	10.30
80165	34.00	80165	31.78
80315	28.00	80315	26.18
80503	37.00	80503	21.58
80629	16.00	80629	13.25
82120	15.00	82120	10.36

Table 6-14. Average LOS for pressure ulcer and non-pressure ulcer cohort

According to Healthcare Financial Management (www.hfm.org), the average cost of hospital stay (years 2003-2008), is \$2000 for Medicare patients and \$2500 for Non-Medicare patients. Also other costs of treating pressure ulcer such as dressings, ointments and specialty beds are estimated to be around \$300, per-day.

2. Increase in LOS Due to Line Infection

A few studies could be found in the literature that have focused on estimating the increase in the LOS, and associated costs due to hospital acquired line infection. Some of these research studies have concentrated on specific hospital population such as low birth weight infants (Payne, et al. 2004) and intensive care unit patients (Digiovine, et al. 1999 and Gracia-Garmendia et al. 1999), for instance. There are also studies that have focused on the general population of the hospital. Battista Orsi, Di Stefano and Noah (2002), estimate the increase in the LOS due to hospital acquired line infection to have a mean of 19.9 days and median of 15 days. Studies also suggest that estimated additional cost per patient due to treatment (replacement of the central venous line (CVL) (approximately \$200), X –Ray and drug administration (approximately \$500, antibiotic costs (between \$100 and \$250 per day) to be close to \$3500, plus the cost of the hospital stay. Kim et al. (2011), Battista Orsi, Di Stefano and Noah (2002) and Digiovine, et al. (1999), have estimation that are relatively consistent with \$3500 cost of treatment.

6.3 Dynamic Model's Validation

Model validation is an essential aspect of any model building methodology in general, and system dynamics modeling in particular. It is a process that involves both formal (quantitative) and informal (qualitative) tools. A model is a simplification of real world to serve a useful purpose and helps us understand a problem/situation. Hence, it has to be determined whether it is good enough for its purpose. The process involves two aspects; first verification, which means ensuring that the equations are technically correct (debugging), and second, validation, which means ensuring that the structure of the model and the assumptions made meet the purpose that the model is intended to serve (Coyle and Exelby, 2000). It is worth mentioning that a valid model is naturally verified; however verification does not guarantee the validity of a model. As Coyle (1977) puts it, validation is “the process by which we establish sufficient confidence in a model to be prepared to use it for some particular purpose”. This confidence building process is a gradual process that is embedded throughout the methodology and starts in the stage of model conceptualization and development and continues even after the implementation of policy recommendations made as the result of the model output (Barlas, 1994, Forrester and Senge, 1980). Although model validation takes places in all the stages of modeling, most of formal (quantitative) validation activities are performed after the model has been constructed.

There are two schools of thought in viewing model validity; first, the empiricist philosophy, which sees a valid model as an objective representation of a real system. In this view models are either correct or incorrect and empirical facts would reveal its truth

or falsehood. In this philosophy, validity is a matter of accuracy rather than usefulness. Second, the more recent relativist school of thought sees a valid model as one of the many possible ways to describe and represent a real situation and believes while no model representation is superior to another in an absolute sense, some could be proven to be more effective and generally models lie in a spectrum of usefulness. System dynamics model validation literature seems agree with the relativist approach to model validity. Hence, besides formal (quantitative) and objective validation of the model, subjective, qualitative and informal components must be involved in validation process to determine the usefulness of a model with respect to a particular purpose (Barlas, 1994).

6.3.1 Informal / Qualitative Model Validation

The informal/ qualitative model validation has been built in the process of model development. The first draft of the model, and the corresponding hypothesis that the financial wellbeing of a healthcare organization i.e. a hospital influences the managerial decisions to reduce costs (e.g. optimizing the length of stay, controlling the operational costs) and expenditure on proactive safety interventions, which in turn effect the risk of experiencing specific adverse events, and the change in the risk of experiencing such adverse events influences the financial standing of the hospital that had originated such dynamics in the system, was discussed and validated with the experts in multiple interview sessions. After gathering domain experts' input on the hypothesis under study in general, and the key players and important factors and relations and feedbacks in particular, the next round of interview discussed the updated version of the model that incorporated all the inputs/suggestions/modifications that the experts made to the previous draft of the model. This procedure was followed in 3 rounds of face-to-face

interview with each expert until a consensus was reached on the current version of the model presented in this thesis. Each expert was then asked to rate the model in terms of completeness, accuracy, ease of understanding and perceived predictability. Experts provided their assessment of the qualitative representation of the model through the following questions:

1. *Completeness*. From your perspective, to what extent does this model capture all important and relevant phenomena for the particular problem that we are studying? On a scale from 0 to 100, 0 would correspond to a model that does not include some important and relevant details, whereas 100 would correspond to a model that includes all details that you consider important. What number would you assign?
2. *Accuracy*: From your perspective, how accurately or realistically does the model depict important feedback effects, and causal chains that influence risk of experiencing adverse events? On a scale from 0 to 100, 0 would correspond to a model that is unrealistic, over-idealized or inaccurate, whereas 100 would correspond to a model that is realistic and accurate. What number would you assign?
3. *Ease of understanding*: From your perspective, how easy is it to understand the overall logic of the model. On a scale from 0 to 100, 0 would correspond to a model that is difficult to follow, even with extensive explanation, and a 100

would correspond to a model that is readily understandable. What number would you assign?

4. *Perceived predictive validity*: From your perspective, if you were to use this model, how well could you predict the change in the risk of specific adverse events as a function of the organizational factors/decisions that influence risk of AEs? On a scale from 0 to 100, 0 would correspond to a model that does not help at all with predicting effects, and a 100 would correspond to a model that predicts the effects very well. What number would you assign?

Table 6-15, reflects the summary of experts' assessment of the system dynamics model.

	Completeness	Accuracy	Ease of Understanding	Predictive Validity
Expert 1	90	90	95	85-90
Expert 2	90	90	90	75
Expert 3	85	85	90	80
Expert 4	90-95	90-95	90-95	85-90
Expert 5	70-80	70-80	80-90	70-80
Expert 6	85	85	85	85
Expert 7	>90	>90	60	Need to see results
Expert 8	100	100	75-80	50
Expert 9	90	90	90	Estimates could be good
Expert 10	90	90	100	100

Table 6-15. Expert opinion; qualitative evaluation of system dynamics module (organizational level contributions to risk)

The procedure that was followed for peer review of Bayesian network models, discussed in section 4.1.3 was also observed here in informal/qualitative validation of the system dynamics model as well. The interview guide that was used to develop/validate and evaluate the model could be found in appendix C.

6.3.2 Formal/ Quantitative Model Validation

The ultimate objective in system dynamics model validation is to establish that the structure of the model is valid. Although we will evaluate the accuracy of the model behavior through certain tests, but this is only meaningful if we are sufficiently confident in the structure of the model. Logically the validation process starts with testing the validity of the structure and follows by evaluating the accuracy of the behavior of the model (Barlas, 1994). Barlas (1996) provides a framework for such a sequence of formal/quantitative validation, and we generally follow this framework for validating the presented model in this study. Figure 6-26, depicts the essence of this framework.

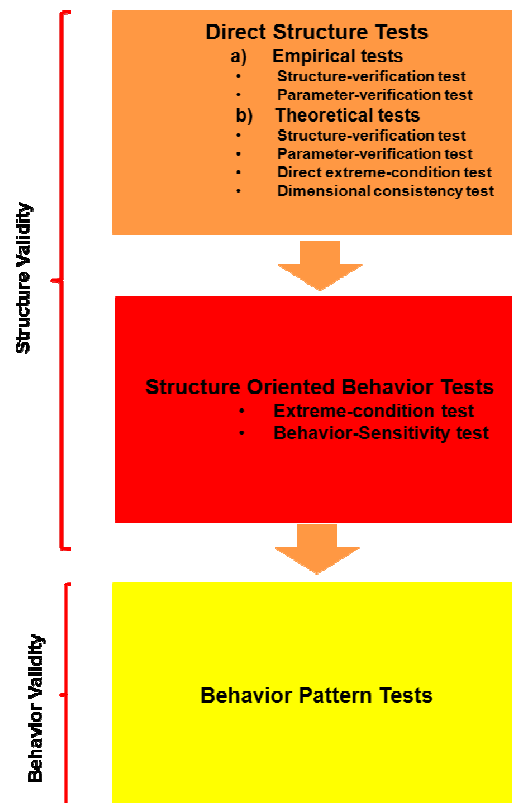


Figure 6-26. System dynamics model validation framework

6.3.2.1 *Direct Structure Tests*

Direct structure tests assess the validity of the model equations individually, and compare them directly to the available knowledge. These tests could be done in two forms;

- Empirically
- Theoretically

The empirical structure test compares model equations against data available from the real system being modeled. Theoretical structure test involves comparing the model equations to the general knowledge on the system found in the literature and/or from domain experts. Forrester and Senge (1980), list Structure and Parameter verification test (comparing the structure and constant parameters of the model against the knowledge of the system conceptually (literature/experts) and numerically (data)) and dimensional consistency test (dimensional analysis of model equations to ensure that the dimensions/units of the equations and parameters are consistent).

- 1) Dimensional Consistency: After careful review of the units used for model parameters and equations, Vensim's "Units Check" (from the menu select: Model>Units check) feature was used to ensure dimensional consistency, and no inconsistencies were found.
- 2) Structure and Parameter Verification: The structure and parameter verification in the case of our mode, involves a combination of empirical and theoretical approach. Some of the equations of the model namely "pressure to Close Revenue Gap", "Pressure to Optimize LOS", "Pressure to Reduce Operational Costs" and "Willingness to Invest in Proactive Safety Interventions" have been validated with

experts, since no empirical information is available for these equations. On the other hand, the rest of the equations in the model, including Probability of Understaffing, Probability of LOS Too Short to Meet Needs”, etc., have been obtained from actual data and other empirical studies available. Section 4.3.2 has detailed discussion on how the equations of the model are obtained.

6.3.2.2 *Structure-Oriented Behavior Tests*

This general category of structure testing indirectly assesses the validity of the structure by applying certain behavior tests. These tests are strong behavior tests that could help the analyst discover possible structural flows. One type of such tests is the Extreme Condition test, which assigns extreme values to selected parameters of the model and compares the behavior generated by the model to the behavior that is expected or observed of the real system being modeled under the same extreme conditions (Barlas, 1994).

A number of these tests have been performed on the model. As an example of extreme condition testing, we assign a value of 1 (maximum) pressure to close revenue gap, which should drive the pressure to optimize LOS and pressure to cut operational cost to the maximum limit (1), and willingness/ability to invest in proactive safety interventions to minimum (0). It also increases the risk of adverse event (in this case risk of pressure ulcer) by 80%. It's worth mentioning that the risk doesn't exponentially increase because at the end of each year, we set the value of pressure ulcer risk to its baseline value

coming from the Bayesian network. Figure 6-27 and Figure 6-28 show the results of these tests.

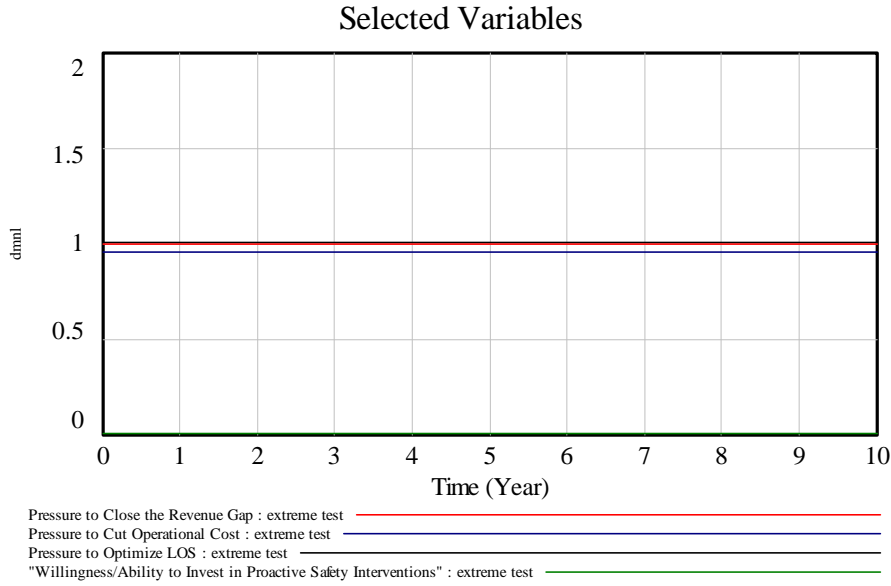


Figure 6-27. Extreme case testing; pressure variables in the model when pressure to close revenue gap is at maximum

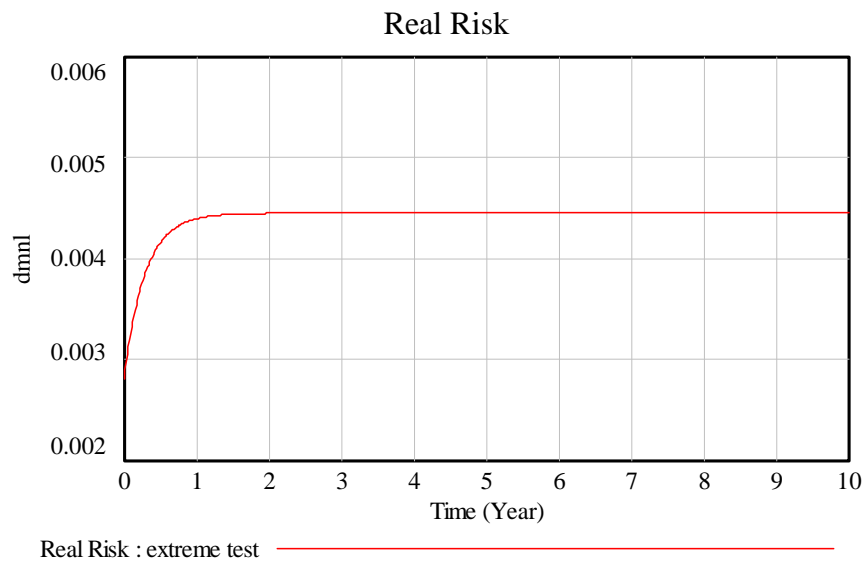


Figure 6-28. Extreme case testing; risk of pressure ulcer when pressure to close revenue gap is at maximum increases 80%

6.3.2.3 Behavior Pattern Test

The two categories of tests above, direct structure test and structure-oriented behavior test are designed to assess the validity of the structure of the model. After building adequate confidence in the structure, we can apply a number of tests that are designed to evaluate the accuracy of the model in reproducing the major behavior patterns of the real system being modeled. Many types of behavior pattern tests could be found in the literature. Forrester and Senge (1980) discuss a number of these tests, including Behavior Reproduction test, Behavior Predictions test, Behavior Anomaly test and Surprise Behavior test among others. The test we have used to evaluate the accuracy of model output is a Frequency Generation test, which falls into the category of Behavior Reproduction tests. The goal here is to see how well the model reproduces the patterns/values that we have observed in the real system. The general idea is to:

1. Calibrate the model based on data available for years 2003-2007
2. Use the model to predict the risk of adverse event (pressure ulcer and line infection) for years 2008-2010
3. Evaluate the accuracy of model prediction with point by point comparison
4. Compare the pattern produced by the model for all the years (2003-2010) with the pattern observed from real data

We should mention that, in the literature on modeling and simulation, a wide range of tests could be found that are based on point-by-point comparison of observed behavior and behavior produced by the model, but these tests are generally less appropriate for system dynamics models. This is because system dynamics models

are usually long term and policy oriented. (Forrester and Segne, 1980, Barlas, 1996). Never the less with limited financial data and reliable adverse event data available (only seven years), this test would build confidence in the accuracy of the model, as well as the structural validity.

We have performed the steps 1 through 4 above, in section 7.1, and have evaluated the performance of the hybrid model, comparing model projections with actual data.

7 The Hybrid Model; Analysis and Results

Section 3.3, explains how the hybrid model which consists of a system dynamics module (to represent the contributions of organizational and policy level factors to the risk of adverse events and the feedback effects of these policies) and a Bayesian Belief Network module (to represent individual patient level and patient provider level factors' contribution to the risk of certain adverse events), functions.

The basic idea is to understand what the base line risk is for any patient for a certain adverse event depending on his/her physiological conditions and provider's decisions in treating this patient, through a Bayesian network. On the other hand, under the influence of financial wellbeing of the hospital and throughput pressures, certain decisions are taken at the hospital level that would affect this baseline risk level, either positively (risk reduction), or negatively. The combination of the two modeling formalisms will inform the decision maker of the overall risk of specific adverse events to be expected, given the individual patient conditions and the existing levels of financial pressures in the system.

7.1 Evaluation and Validation of the Hybrid Model

Our goal here is to see whether the combination of system dynamics model and the Bayesian belief networks (the hybrid model) can reproduce the patterns and the values or risk of specific adverse events from actual clinical data.

7.1.1 Hybrid Model Performance; Risk of Pressure Ulcer

Step 1: Calibrate the model based on data available for years 2003-2007

Nine parameters (p_1, \dots, p_9), that determine the shape of the pressure functions (Pressure to close the revenue gap, Pressure to optimize LOS, Pressure to cut operational costs and Willingness/Ability to invest proactively in safety interventions) have been calculated the way that best fit experts' input on these pressures (section 6.2.1).

Based on data for 2003-2007, we calibrate/optimize these values, so that the error (i.e. error between model's prediction on risk of pressure ulcer and the actual relative frequency of hospital acquired pressure ulcer obtained from clinical data, for each year) is minimized.

Doing so, the following values are obtained for p_1, \dots, p_9 :

$$p_1 = 0.084431$$

$$p_2 = 75.5669$$

$$p_3 = 0.0439$$

$$p_4 = 1.0771$$

$$p_5 = 0.3406$$

$$p_6 = 1.8830$$

$$p_7 = 0.1073$$

$$p_8 = 0.2$$

$$p_9 = 24.1834$$

The payoff function ($error^2$) from this optimization is 3.55082E-8.

Step 2: Use the model to predict the risk of pressure ulcer for other years

In the hybrid model, baseline risk of pressure ulcer is provided by the pressure ulcer BBN model. However, the system dynamic model and the pressure ulcer BBN, share a same node; Understaffing probability (or as is called in the BBN, probability of staff adequacy). So the following steps are taken in the hybrid model, for calculating the risk of pressure ulcer for each year:

1. Calculate the baseline probability of pressure ulcer from pressure ulcer BBN
2. Input the baseline probability of pressure ulcer to the system dynamics model
3. The system dynamics model, calculates the probability of understaffing (or staff adequacy) for year “i”
4. The pressure ulcer BBN reads this value (probability of staff adequacy) from the system dynamics module, and calculates a new baseline probability of pressure ulcer for year ‘i+1’
5. Go to step 1

Figure 7-1 depicts risk of pressure ulcer for different years projected by the hybrid model, versus the actual clinical data for relative frequency of hospital acquired pressure ulcer. Table 7-1 shows the error of hybrid model’s predictions for risk of pressure ulcer, compared to the actual data.

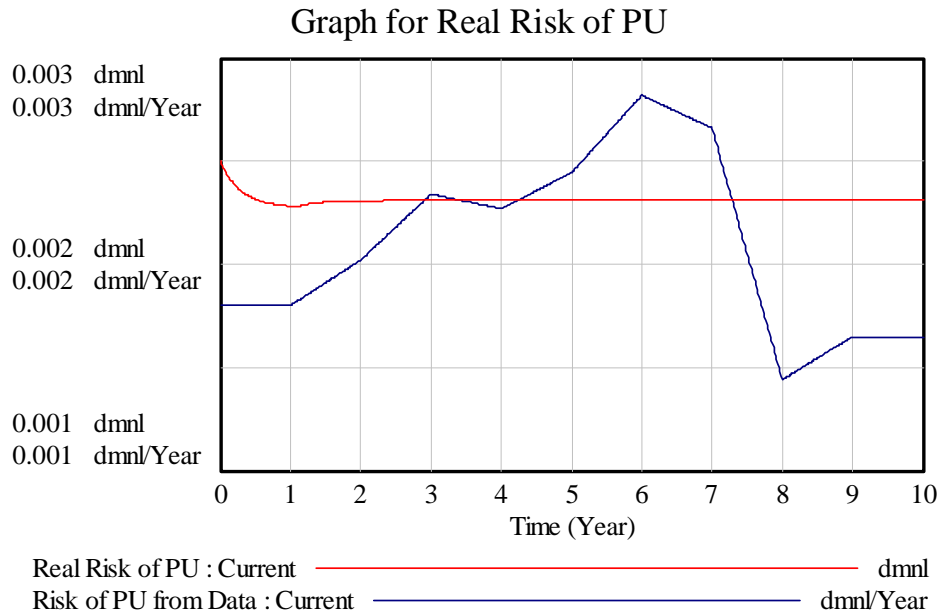


Figure 7-1. Risk of hospital acquired pressure ulcer; hybrid model predictions for each year versus actual data

Year	Baseline Risk of Pressure Ulcer	Hybrid Model :Risk of Pressure Ulcer	Data: Actual Risk of Pressure Ulcer	Error
2003	0.0024 (from BBN)	0.0022	0.00181	18%
2004	0.0025	0.0021	0.00202	4%
2005	0.0025	0.0023	0.00234	2%
2006	0.0025	0.0022	0.00227	3%
2007	0.0025	0.0022	0.00244	11%
2008	0.0025	0.0022	0.00282	28%
2009	0.0025	0.0022	0.00266	21%
2010	0.0025	0.0022	0.00144	35%
2011	0.0025	0.0022	0.00165	25%

Average Error	15%
----------------------	------------

Table 7-1. Hybrid models' error in predicting the probability of pressure ulcer for each year

Note that the baseline risk of pressure ulcer, remains the same from 2004-2011. This is due to the fact that the input from system dynamic model to pressure ulcer BBN, the probability of understaffing (i.e. probability of staff adequacy), remains the same at 0.1. This is expected, because the particular institution that we have gathered our clinical data

from is financially well and the pressure to close financial gap is very low throughout these years.

7.1.2 Hybrid Model Performance; Risk of Line Infection

For the hybrid model, consisting of the system dynamics module and line infection BBN, we will follow the same steps as section 7.1.1.

Step 1: Calibrate the model based on data available for years 2002-2006

Nine parameters (p_1, \dots, p_9), that determine the shape of the pressure functions have been calculated the way that best fit experts' input on these pressures (section 6.2.1).

Based on data for 2002-2006, we calibrate/optimize these values, so that the error (i.e. error between model's prediction on risk of pressure ulcer and the actual relative frequency of hospital acquired pressure ulcer obtained from clinical data, for each year).

Doing so, the following values are obtained for p_1, \dots, p_9 :

$$p_1 = 0.094$$

$$p_2 = 76.30$$

$$p_3 = 0.049$$

$$p_4 = 2.00$$

$$p_5 = 0.580$$

$$p_6 = 0.960$$

$$p_7 = 0.107$$

$$p_8 = 1.770$$

$$p_9 = 26.70$$

The payoff function ($error^2$) from this optimization is 2.8E-5.

Step 2: Use the model to predict the risk of line infection for other years

In the hybrid model, baseline risk of pressure ulcer is provided by line infection BBN model. However, the system dynamic model and the line infection BBN, share a same node; Understaffing probability (or as is called in the BBN, probability of staff adequacy). So the following steps are taken in the hybrid model, for calculating the risk of line infection for each year:

1. Calculate the baseline probability of line infection from line infection BNN
2. Input the baseline probability of line infection to the system dynamics model
3. System dynamics model, calculates the probability of understaffing (or staff adequacy) for year “i”
4. Line infection BBN reads this value (probability of staff adequacy) from system dynamics module, and calculates a new baseline probability of line infection for year ‘i+1”
5. Go to step 1

Figure 7-2, depicts risk of line infection for different years projected by the hybrid model, versus the actual clinical data for relative frequency of line infection. Table 7-2, shows the error of hybrid model’s predictions for risk of line infection, compared to the actual data.

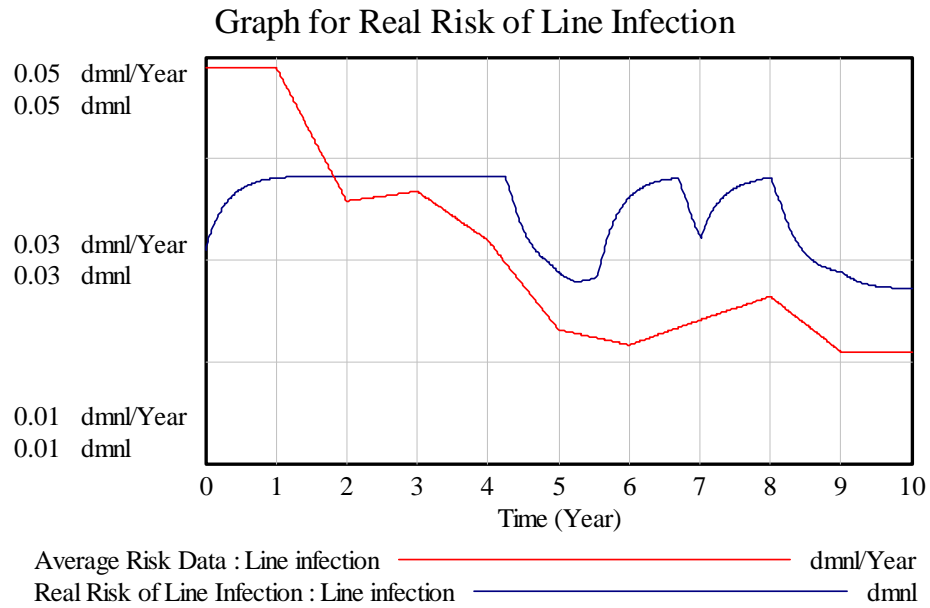


Figure 7-2. Risk of line infection; hybrid model predictions for each year versus actual data

Year	Baseline Risk of Line Infection	Hybrid Model :Risk of Line Infection	Data: Actual Risk of Line Infection	Error
2002	0.0302 (from BBN)	0.038	0.0498	31%
2003	0.031	0.0382	0.0385	1%
2004	0.031	0.0382	0.03677	4%
2005	0.031	0.0382	0.03195	16%
2006	0.031	0.0288	0.02315	20%
2007	0.031	0.0361	0.02164	40%
2008	0.031	0.0323	0.0241	25%
2009	0.031	0.0271	0.02	26%
Average Error				20%

Table 7-2. Hybrid models' error in predicting the probability of line infection for each year

7.1.3 Hybrid Model Performance; Risk of Line Infection and Pressure Ulcer

One can argue that since the system dynamic model is capturing the organizational level contributions to risk, it affects the risk of pressure ulcer and line infection, both at the same time and if the model is performing correctly, it should capture these effects on risk of both of the adverse events simultaneously. In sections 7.1.1 and 7.1.2, we evaluated the effect of decision/policy level factors on each of the adverse events we have modeled (pressure ulcer and line infection) separately. Here we will evaluate the performance of the model where organizational factors/decisions affect risk of these adverse events simultaneously. In other words we will quantify the hybrid model that consists of 3 models: pressure ulcer BBN, line infection BBN and the system dynamics module for organizational effect.

For the hybrid model, consisting of the system dynamics module and both pressure ulcer and line infection BBNs, we will follow the same steps as described in section 7.1.1.

Step 1: Calibrate the model based on data available for years 2003-2006

Nine parameters (p_1, \dots, p_9), that determine the shape of the pressure functions have been calculated the way that best fit experts' input on these pressures (section 6.2.1).

Based on data for 2003-2006, we calibrate/optimize these values, to minimize error (difference between model prediction and actual data). In this case though our error term consists of error in predication for line infection risk plus error of prediction for pressure ulcer risk, doing so, the following values are obtained for p_1, \dots, p_9 :

$$p_1 = 0.11$$

$$p_2 = 67.38$$

$$p_3 = 0.15$$

$$p_4 = 1.22$$

$$p_5 = 0.47$$

$$p_6 = 0.008$$

$$p_7 = 0.1073$$

$$p_8 = 0.16$$

$$p_9 = 29.3$$

The payoff function ($error^2$) from this optimization is 1.38E-5.

Step 2: Use the model to predict the risk of line infection and pressure ulcer for other years

In the hybrid model, baseline risks of both adverse events are provided by the BBN models. Since probability of staff adequacy (i.e. understaffing probability in system dynamics model) is shared between both BBNs and the dynamic model, the following steps are taken, for calculating the risk of both adverse events for each year:

1. Calculate the baseline probability of both adverse events from their respective BBN
2. Input the baseline probability of line infection to the system dynamics model
3. System dynamics model, calculates the probability of understaffing (or staff adequacy) for year “i”
4. The BBNs read this value (probability of staff adequacy) from system dynamics module, and calculates a new baseline probability of adverse event for year ‘i+1”
5. Go to step 1

Figure 7-3 and Figure 7-4, depict risks of pressure ulcer and line infection for different years projected by the hybrid model, versus the actual clinical data for relative frequency of these adverse events. Table 7-3, shows the error of hybrid model's predictions for the risks of these adverse events, compared to the actual data.

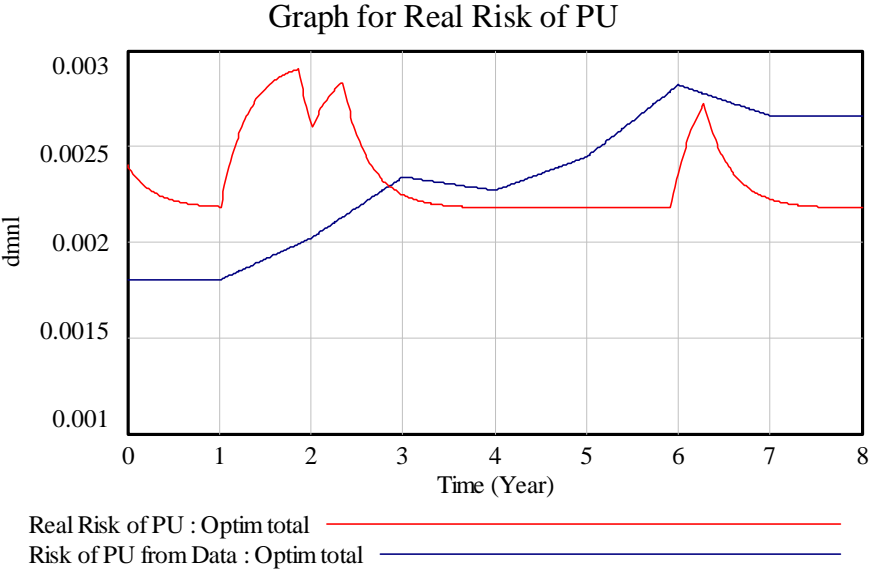


Figure 7-3. Risk of PU; hybrid model predictions for each year versus actual data calibrating the model using both BBNs

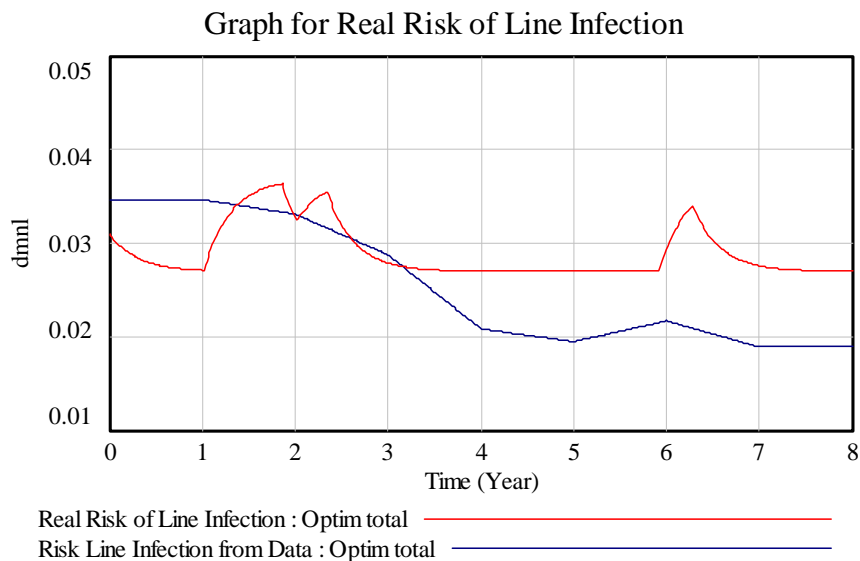


Figure 7-4. Risk of line infection; hybrid model predictions for each year versus actual data calibrating the model using both BBNs

Year	Hybrid: Risk of Pressure Ulcer	Actual Risk of Pressure Ulcer	Error	Hybrid: Risk of Line Infection	Actual Risk of Line Infection	Error	
2003	0.0021	0.00181	13.81%	0.0271	0.0385	42%	
2004	0.0026	0.00202	22.31%	0.0327	0.03677	12%	
2005	0.0022	0.00234	6.36%	0.0278	0.03195	15%	
2006	0.0021	0.00227	8.10%	0.027	0.02315	14%	
2007	0.0021	0.00244	16.19%	0.0269	0.02164	20%	
2008	0.0028	0.00282	0.71%	0.0293	0.0241	18%	
2009	0.0022	0.00266	20.91%	0.0275	0.02	27%	
Average Error			12.63%	Average Error			21%

Table 7-3. Hybrid models' error in predicting the probability of line infection for each year, calibrating with both BBNs

7.2 Use of the Risk Model; Hypothetical Examples

A. Patient A's risk of line infection

Suppose patient A, is scheduled for a surgery in June 2011, and will need a catheter line. Due to his procedure though, he will have to have the line inserted in femoral vein (versus more desirable subclavian vein). Also assume that the patient due to his clinical condition and age has diminished resistance. What will be his chances of acquiring line infection?

To answer this question and assess patient A's risk of line infection, the following procedure will be followed using the hybrid model:

1) Estimate patient A's baseline probability of line infection

This is done by setting the information we have from patient A, as evidence to the model (Figure 7-5). In other words: Patient's Constraint: T (with probability of 1), and Patient Resistance Factors: Diminished (with probability of 1). Note that we can also set soft evidence for any node, if we are uncertain about the presence of a certain condition.

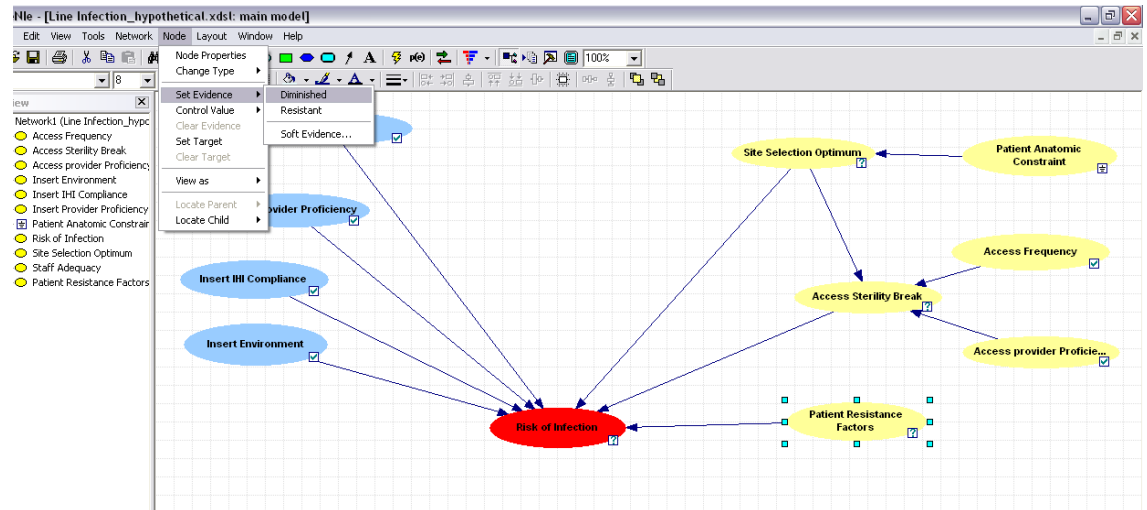


Figure 7-5. Setting evidence in line infection BBN

- 2) Considering that line infection BBN and the system dynamics module have a common node (probability of understaffing or staff adequacy), load this probability from system dynamic model into line infection BBN.
- 3) Update the line infection BBN with this information to get the baseline probability of line infection for patient A (Figure 7-6).
- 4) The baseline probability of line infection for this patient, increases from 0.0302 to 0.0386 (28% increase).

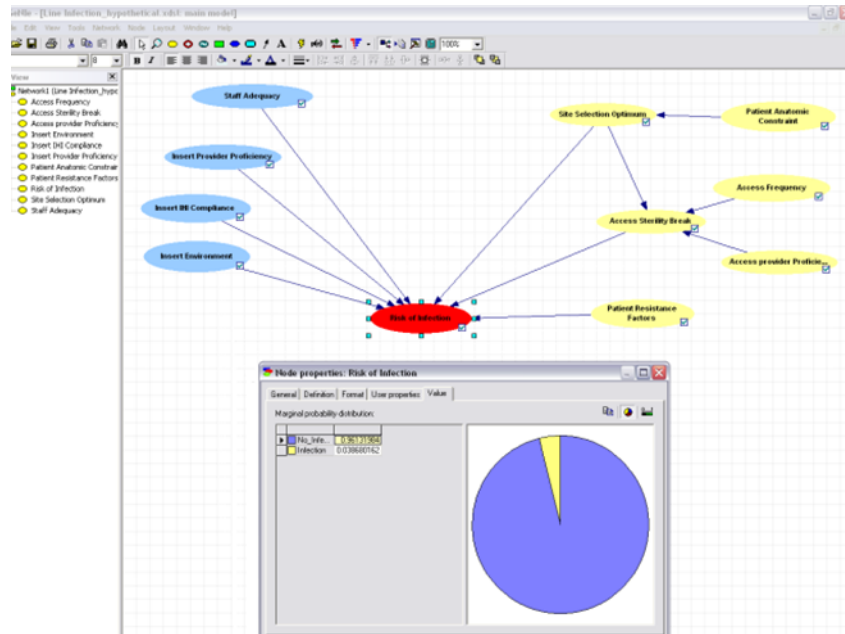


Figure 7-6. Compiling BBN to update the probability of line infection

- 5) Input this baseline probability of line infection, and predict patient A's probability of line infection for 2011, using the hybrid model (Figure 7-7).

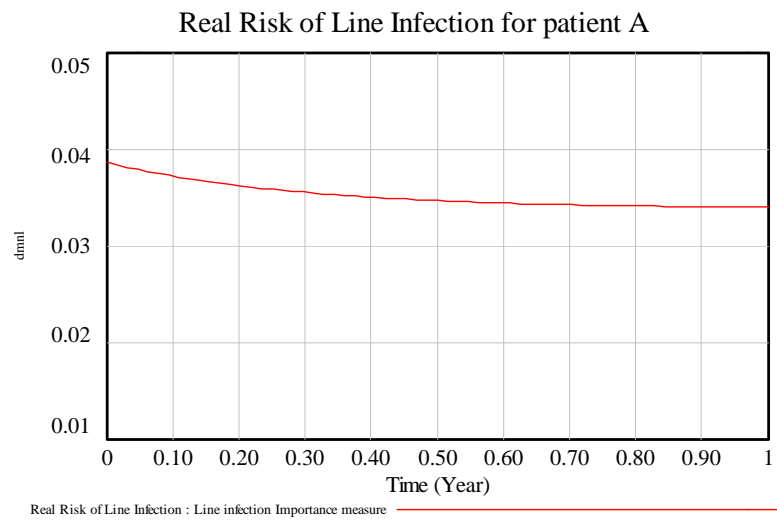


Figure 7-7. Predicting patient A's probability of line infection, influenced by organizational factors

The model projects risk of line infection for patient A to be 0.0339 in 2011, which indicates a 12% decrease. Considering that hospitals operating margin for 2011 so far has been at 3.5%, this reduction is expected since the production pressures are low and ability to invest proactively in safety are high.

B. As another hypothetical example, assume that the hospital at 2011 has operated on a 3.5%, and they set a goal of increasing their operating margin 0.2%, each year. The hybrid model can predict that in the next 10 years, the baseline risk of line infection could be reduced by 13%, from 3.05% to 2.66%. This implies over 2.7 million dollars in savings each year.

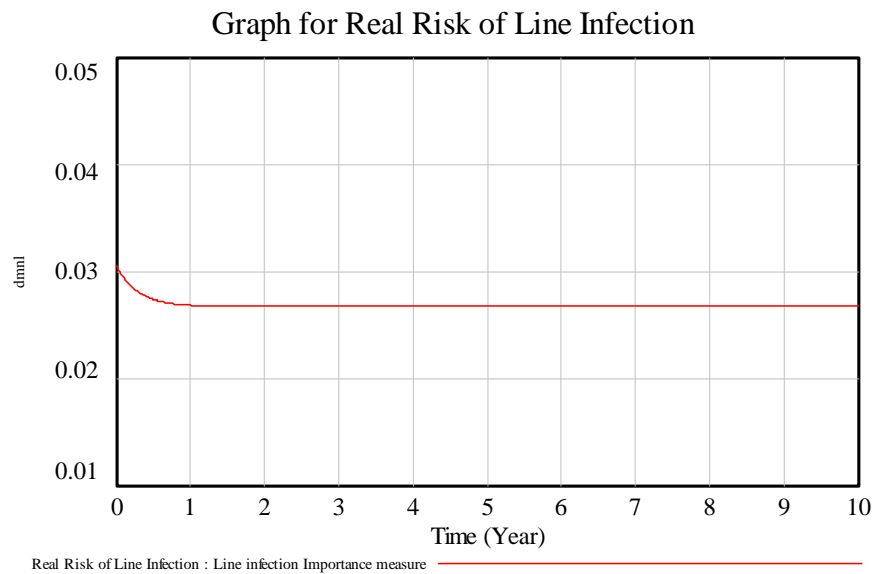


Figure 7-8. Decline in the risk of line infection over the next 10 years, due to increase in operating margin by 0.2% per year

Similarly, if hospital at 2011 has operated on a -1%, and their operating margin declines 0.2% every year, over the next 10 years, the baseline risk of line infection will increase by 23%, from 3.05% to 3.76%, which implies an expected over 6 million dollars in expenses

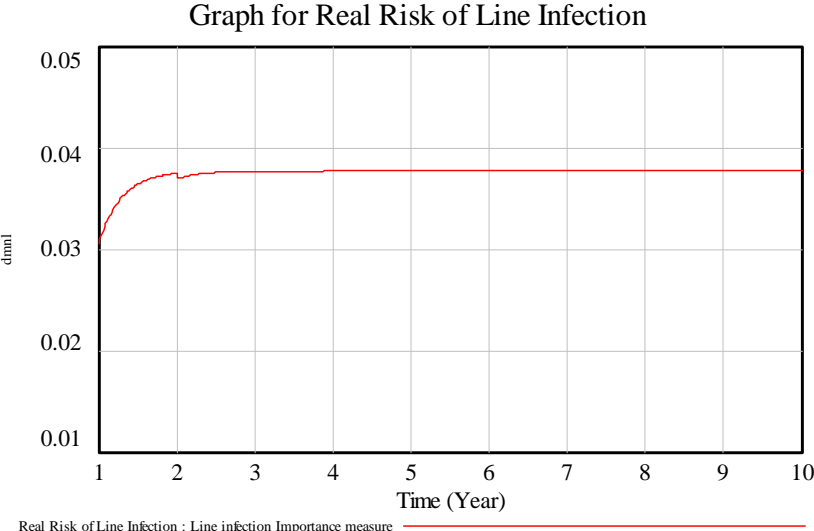


Figure 7-9. Increase in the risk of line infection over the next 10 years, due to decrease in operating margin by 0.2% per year

7.3 Uncertainty Analysis

7.3.1 Uncertainty Analysis; Hybrid Model for Risk of Pressure Ulcer

As a measure of goodness of an estimate, to examine how closely the estimated values relate to reality and as a basis for decision making we can perform a set of uncertainty analyses (Modarres, 2006). These uncertainty analyses, show the impacts of analyst's assumptions, variability in the parameters, impact of data incompleteness, and the effect of expert opinion. These uncertainties are represented by probability distributions which are then propagated through the entire risk model (Smith, 2011).

When estimating the parameters of the Bayesian networks and the system dynamics module, throughout chapter 4 of this dissertation, we have represented the uncertainties with the appropriate probability distributions. In discussing the hybrid model's performance in previous section, 7.1, we have used the mean values of those distributions. In this section, we will study the effects of parameter uncertainties using their respective probability distributions, on the hybrid model as a whole. Many univariate or multi-variate uncertainty analyses could be performed to propagate parameter uncertainty over the hybrid model. Below are two examples of these uncertainty analyses runs.

a) Uni-variate: baseline risk of pressure ulcer uncertainty; Uniform (0.001,0.003)

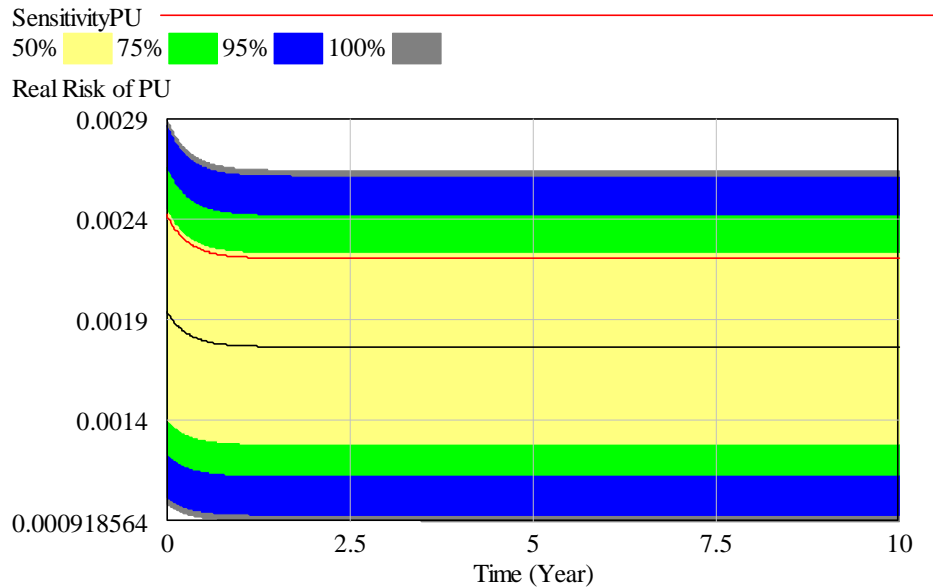


Figure 7-10. Confidence bounds for model predictions on risk of pressure ulcer as a result of uncertainty over baseline risk of pressure ulcer

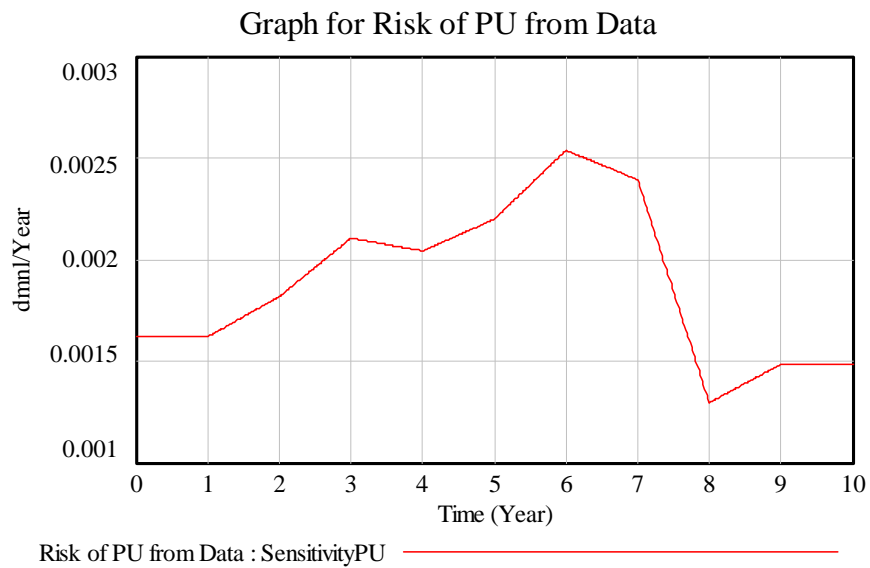


Figure 7-11. Risk of pressure ulcer from clinical data

b) Multi-variate: probability of prolonged LOS and ; Normal (0.045,0.38) and baseline risk of pressure ulcer uncertainty; Uniform (0.001,0.003)

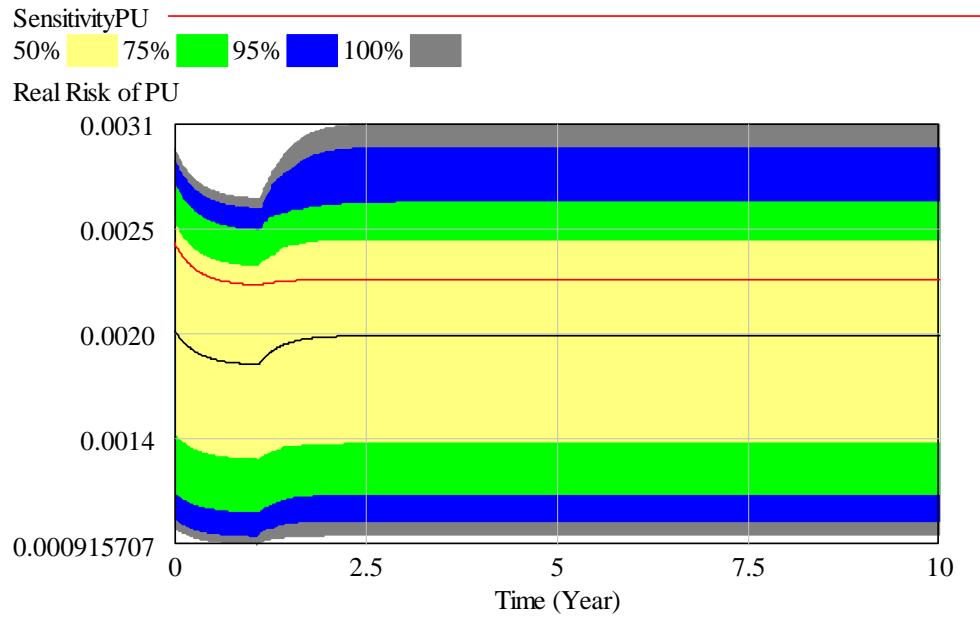


Figure 7-12. Confidence bounds for model predictions on risk of pressure ulcer as a result of uncertainty over baseline risk of pressure ulcer and uncertainty over the probability of prolonged LOS

7.3.2 Uncertainty Analysis; hybrid Model for Risk of Line Infection

Similar to the case of hybrid model for risk of pressure ulcer, many uni-variate or multi-variate uncertainty analyses could be performed to propagate parameter uncertainty over the hybrid model. Below are two examples of these uncertainty analyses runs.

- a) Uni-variate: baseline risk of line infection uncertainty; Weibull distribution
 $(\alpha = 3.221, \beta = 0.0312)$

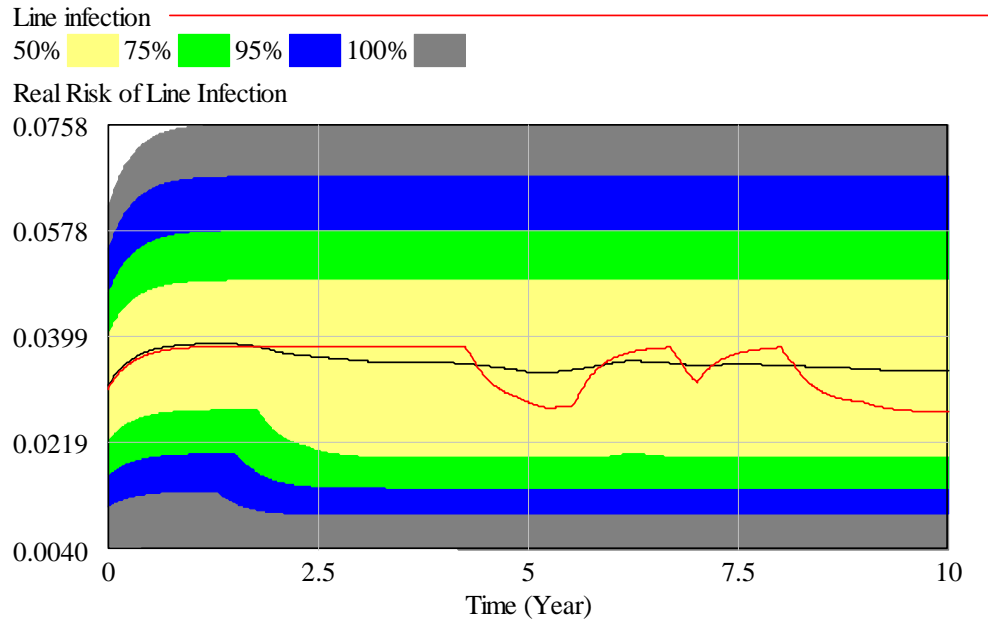


Figure 7-13. Confidence bounds for model predictions on risk of line infection as a result of uncertainty over baseline risk of line infection

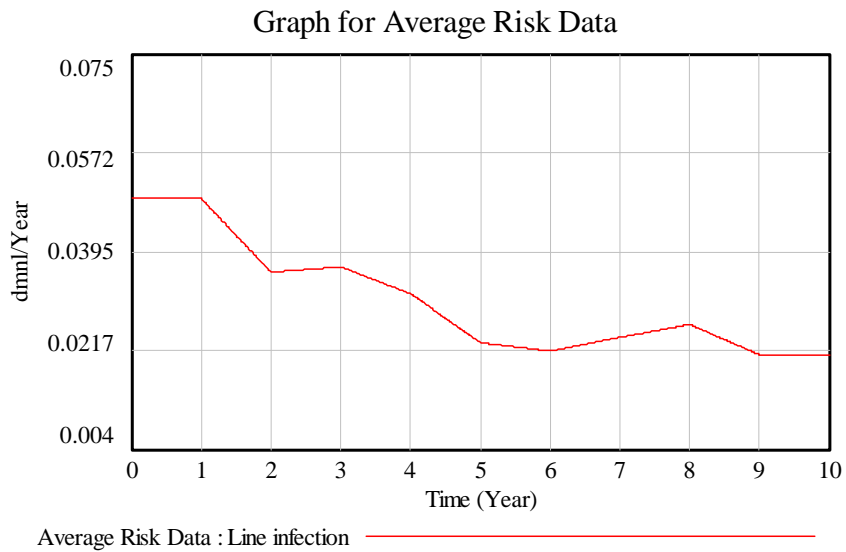


Figure 7-14. Risk of line infection from clinical data

- a) Multi-variate: probability of prolonged LOS and ; Normal (0.045,0.38) and baseline risk of pressure ulcer uncertainty; Weibull distribution ($\alpha = 3.2206, \beta = 0.03119$)

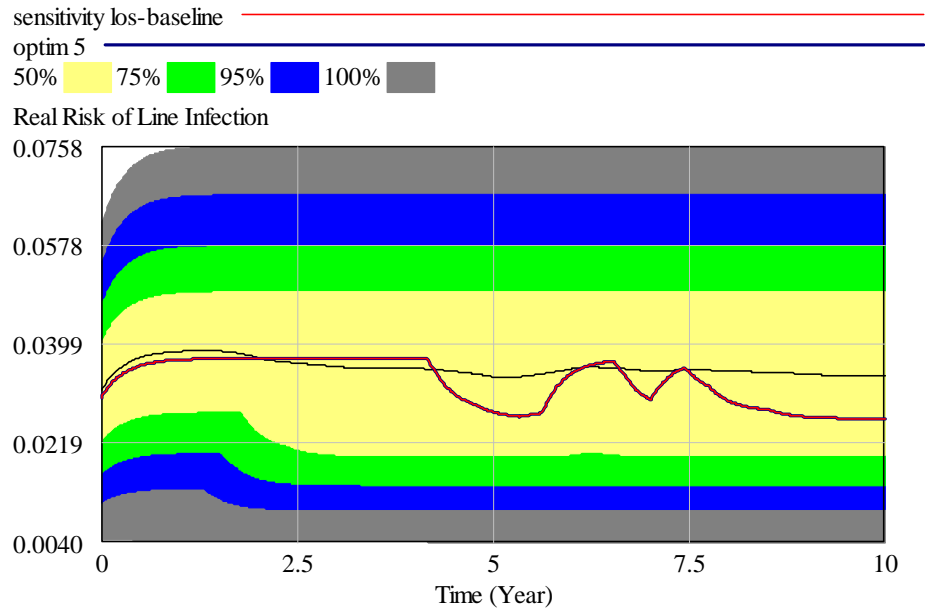


Figure 7-15. Confidence bounds for model predictions on risk of line infection as a result of uncertainty over baseline risk of pressure ulcer and uncertainty over the probability of prolonged LOS

7.3.3 Uncertainty Analysis; hybrid Model for Risk of Line Infection and Pressure Ulcer Combined

Similar to what we discussed in sections 7.3.1 and 7.3.2, many uni-variate or multi-variate uncertainty analyses could be performed to propagate parameter uncertainty over the hybrid model that consists of the dynamic model and two BBNs for pressure ulcer and line infection risks. Below is an example of these uncertainty analyses runs.

- a) Uni-variate: baseline risk of line infection uncertainty; Weibull distribution
($\alpha = 3.2206, \beta = 0.03119$)

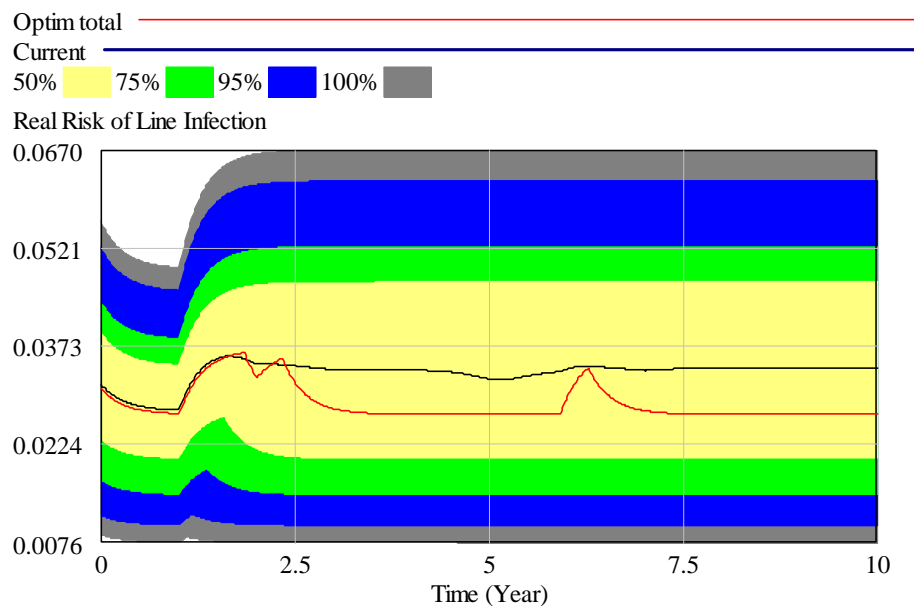


Figure 7-16. Confidence bounds for model predictions on risk of line infection as a result of uncertainty over baseline risk of line infection; the case of the hybrid model with both line infection and pressure ulcer BBNs

7.4 Importance Measures

An important step in any risk analysis activity is to identify the elements of the system that have the most contribution to system risk. The common metrics used in identifying such contributions is the importance ranking. Identification of major risk contributors using importance measures can give direction to risk management activities, and guide allocating resources into areas which will have the highest impact on the system's risk reduction (Modarres, 2006). Birnbaum, Fussel-Vesely (FV), Risk Reduction Worth (RRW), and Risk Achievement Worth (RAW) are among commonly used risk importance measures.

7.4.1 Risk Reduction Worth

The Risk Reduction Worth (RRW) importance, measures change in risk of the system, when a risk element is perfect; i.e. a component's failure probability is assumed to be zero. In other words it measures how much system's risk could be improved if one event could be fixed, and shows, theoretically, what is the limit of the performance improvement of the system (Modarres, 2006).

$$IM_{RRW} = \frac{R}{R(P_i = 0)}$$

Where R is the total system risk, and $R(P_i = 0)$, is system risk when risk element "i" is made perfect. For instance, some of the risk elements in the case of this study could be decisions made at the organizational level to address financial deficiencies in the form of

“Pressure to Optimize LOS” (P_1), “Pressure to Cut Operational Costs” (P_2) and “Willingness/Ability to Invest in Proactive Safety Interventions” (P_3).

Since the hybrid model, has a dynamic element built into it (the system dynamics module), which captures how decisions and policies that contribute to the risk of specific adverse events evolve over time, using risk importance measures we can also project how the importance of these decisions in system risk, may evolve and change over time.

7.4.2 Dependencies in Risk Importance Measure’s Quantification

Due to the fact that the modules in the hybrid methodology; the system dynamic module and the Bayesian network, may share one or more nodes (i.e. Staff Adequacy in the model we have developed here), can cause dependencies. To eliminate this dependency, the following procedure needs to be followed.

1. Find the risk of specific adverse event (i.e. R) using the hybrid model for a specific year.
2. If the event of the interest (risk element), P_i is in the Bayesian belief network, assume perfect condition for P_i ; propagate the Bayesian network one time and store the intermediate probability of adverse event to be used by the second layer of the hybrid model (i.e. the system dynamics module), while still assuming perfect condition for P_i in the system dynamics module, and quantify the hybrid model, which will project how system risk will evolve over time assuming perfect condition for risk element P_i .

3. If the event of the interest (risk element), P_i is not in the Bayesian belief network, find $R(P_i = 0)$, by quantifying the hybrid model again

4. The risk reduction worth measure of the event of interest is obtained from:

$$\frac{R}{R(P_i = 0)}$$

5. Assuming decision $P_i = 0$ is taken, which means :

- a. $P_1 = 0$ relaxing pressure to optimize LOS
- b. $P_2 = 0$ relaxing pressure to cut operational costs
- c. $P_3 = 0$ relaxing safety investment constraints

This procedure is inspired by the procedure Wang (2007) has proposed for calculating importance measures in static models, consisting of ESDs, FTs and BBNs.

7.4.3 Example; Importance Measure for Pressure to Optimize LOS in the Risk of Line Infection

To obtain importance measure for pressure to optimize LOS, for example, we need to calculate the risk of line infection for each year, when the pressure to optimize LOS is at minimum 0. We also need to calculate risk of line infection for each year without interfering with model variables, which we have done and the results are reflected in Table 7-2. We do this for pressure to cut operational cost and willingness/ability to invest proactively in safety. Table 7-4 shows the importance measure for these factors/pressures, for each year of analysis.

Year	Hybrid Model :Risk of Line Infection	Risk of Line Infection; pressure to optimize LOS=0	Importance Measure
2002	0.037	0.0382	0.968586387
2003	0.0372	0.0385	0.966233766
2004	0.0275	0.0384	0.716145833
2005	0.0275	0.0384	0.716145833
2006	0.0343	0.029	1.182758621
2007	0.0295	0.0364	0.81043956
2008	0.0265	0.0289	0.916955017
2009	0.0264	0.0272	0.970588235
Average importance over 9 year period			0.905981657

Year	Hybrid Model :Risk of Line Infection	Risk of Line Infection; pressure to cut operational costs=0	Importance Measure
2002	0.037	0.0367	1.008174387
2003	0.0372	0.0368	1.010869565
2004	0.0275	0.0369	0.745257453
2005	0.0275	0.0369	0.745257453
2006	0.0343	0.0267	1.284644195
2007	0.0295	0.0337	0.87537092
2008	0.0265	0.0291	0.910652921
2009	0.0264	0.0271	0.974169742
Average importance over 9 year period			0.944299579

Year	Hybrid Model :Risk of Line Infection	Risk of Line Infection;Willingness to invest in safety=1	Importance Measure
2002	0.037	0.0284	1.302816901
2003	0.0372	0.0283	1.314487633
2004	0.0275	0.0274	1.003649635
2005	0.0275	0.0271	1.014760148
2006	0.0343	0.0271	1.265682657
2007	0.0295	0.0281	1.049822064
2008	0.0265	0.0284	0.933098592
2009	0.0264	0.0271	0.974169742
Average importance over 9 year period			1.107310921

Table 7-4.Importance measures of pressure functions in the model for years 2002-2009

We can compare the importance of these pressures year by year, or we can alternatively compare their average importance over the course of 8 years. The results indicate that investment in proactive safety interventions is the most important decision factor in terms of influencing risk. Operational budget is the second most important influencing factor in risk, and optimizing LOS comes third.

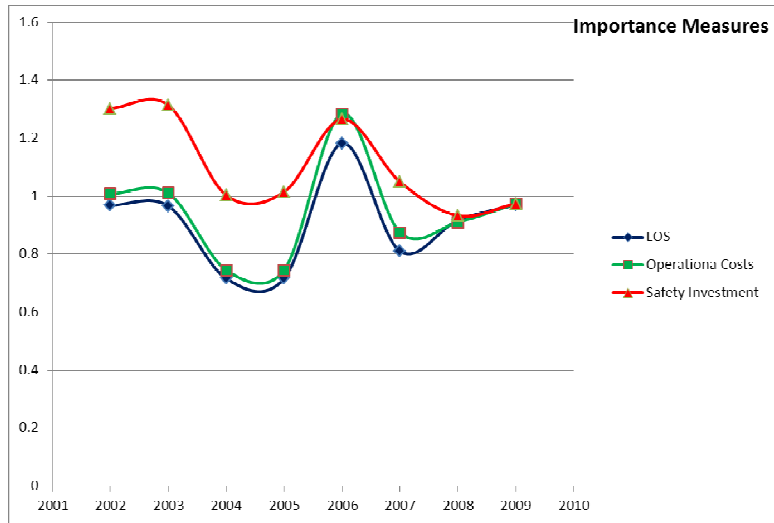


Figure 7-17 .Importance measures of model's pressure functions over time

Figure 7-17 depicts the importance of each of these pressure functions over time.

7.5 Model Requirements for Application

The evaluation and the validation of the individual modules (BBNs and the system dynamic module) and the hybrid model, show that the models developed here have potential to be used as a predictive model for decision making purposes, and to capture the major dynamics of healthcare organizations that have an effect on the risk of adverse events. We have demonstrated this with limited data that was available to us. To further strengthen confidence in the accuracy and predictive power of the model, additional rigorous validation with additional data is required. This involves;

- More expert opinion, from a diverse set of hospitals, on the soft factors in the model
- More adverse event data from a variety of hospitals, although finding clean reliable clinical data, as we have tried to collect and use in the models in this study, could be challenging to say the least
- Meticulous modeling of the cost and reimbursement structure. The data we had available on financial records, consisted of operating margins and total revenue and cost for a few years. Detailed modeling of cost and reimbursement structure will not only increase the accuracy but also make the model a dynamic model in its true sense.
- Modeling more adverse events (in addition to the two BBNs we have developed so far for line infection and pressure ulcer)

In the models we have developed in this study, some factors are hospital-specific factors. For instance hospitals may respond to revenue gap differently. In our interviews with the

experts, they revealed that while hospitals do take one or a combination of three decisions we have modeled here (optimizing LOS, reducing operational costs, and level of investment in safety) to address revenue problems, but the order and intensity in which they implement these decisions may differ from hospital to hospital.

Perhaps, collecting expert opinion from different categories of hospitals and modeling the pressure functions for that specific hospital category, will customize the model for a specific hospital category, and hence make for a better decision making tool.

8 Summary and Conclusion

In this study, we have proposed a hybrid modeling methodology, capable of modeling the risk of hospital acquired adverse events, more realistically. This hybrid modeling environment, consists of Bayesian belief networks and system dynamics modeling formalism. The Bayesian belief networks are used to capture patient level, and patient provider level factors that may affect the risk of a certain adverse event. On the other hand, using system dynamics formalism, we can capture risk contributors at the level of organization, including production pressures, pressure to reduce operational costs; pressures to optimize (minimize) length of stay and pressures that impose limitations on what healthcare organizations can spend on proactive safety interventions. These pressures are mainly imposed upon the system, by financial constraints. On the other hand external agencies and third party payers (e.g. insurance companies) increase this pressure by penalizing the hospital for the occurrence of adverse events.

Employing this methodology, we have developed a dynamic model for system level risk factors, and two Bayesian network models for two specific adverse events; pressure ulcer, and line infection.

These models have been developed using the factors we found in a thorough literature search (believed/proven to be influencing risk factors both at the hospital level model (SD), and specific adverse event models (BBNs)), and expert opinion. A Panel consisting of 17 experts from a number of healthcare organizations, with years of clinical and patient safety experience, was interviewed in person, in multiple sessions (resulting in over 120 hours of interview) in the process of developing and validating these models.

We also used 8 years clinical data from one of Harvard's teaching hospitals, to validate the models, both the BBN level, and the hybrid model.

This new approach provides a more realistic view and captures the dynamics of risk/safety as a function of policy and organizational decisions. The methodology could be used as a tool to predict the unintended consequences of internal and external decisions and policies on safety, and as a tool to investigate the impacts of policy modifications and to optimize decision making. It is also conceivable to use at the level of individual healthcare organization as well as external agencies (e.g. Federal and private insurers).

8.1 Challenges

Aside from the usual administrative challenges in arranging interviews with the domain experts, who typically have extremely busy schedules, perhaps the most important challenge is obtaining reliable and relevant data for developing and validating the models. In validating the models, we have tried to use clinical data as much as possible and elicited experts' opinion, where data was unavailable or unreliable. Not all the factors we have in the model are actually observed and recorded in hospitals (e.g. staff adequacy). As our clinical experts put it, healthcare data is quite "messy".

Especially in the case of the dynamic module, that we have used soft factors to represent system level pressures (e.g. pressure to optimize LOS), we had to solely rely on experts' assessment. We have made effort to calibrate these opinions with the data that was available to us.

8.2 Contributions

The contributions of this research could be summarized as follows in two categories; A) Risk modeling methodology and B) Causal model development

A. Risk modeling methodology

- Selection and integration of suitable methods for modeling risk in healthcare
- Hybrid SD/BBN
- Development of uncertainty/sensitivity analysis procedure for the hybrid methodology
- Development of RIM for the hybrid methodology

B. Causal model development

- Development of dynamic model for organizational level decisions/factors
- Development of BBN causal models for 2 common adverse events
- Collection and analysis of data and expert opinion for model construction and parameter estimation
- Introduction of new parameters to address cause and effect relations among tangible and intangible phenomena
- Use of Bayesian methods for inference with expert opinion
- Use of Bayesian model uncertainty treatment method to improve model calibration and address data gaps

8.3 Future Work

The next logical step for this study would be evaluating and validating the model with more data.

- Data for more number of years
- Clinical data and expert opinion from more experts and variety of hospitals

We have used a 9 years of clinical data to evaluate the performance of the model. We have calibrated 9 parameters (for pressure functions) of these models using these data. One can question the confidence level on this calibration, where 9 degrees of freedom are determined with only 9 years' worth of data. Also, all of the experts that participated in this study and provided their opinion on various aspects of the models, are clinicians and healthcare professionals that are practicing in some the world's best hospitals. We need to incorporate expert opinion and also clinical data from other hospitals that do not necessarily fall into this category.

Another aspect of this model that could be improved is that some of nodes in the model need to be modeled in more detail. More specifically for operating margin, that is basically the driving engine of the hybrid model, we have only used the estimates available to us for operating margins and total cost and reimbursement for a few years. More comprehensive modeling of the cost and reimbursement structure is definitely needed for more accurate results.

Appendix A. Model Validation Interview Guide, Pressure Ulcer BBN

Purpose:

- A. Introducing experts to the concept of Bayesian belief networks, the structure and the concept of conditional probabilities
- B. Introducing the preliminary BBN model for risk of pressure ulcer
- C. Eliciting expert's opinion about the factors and relations in the model, addition or deletion of the factors if necessary based on experts judgment
- D. Eliciting expert's quantitative assessment on some of the parameters of the model
- E. Qualitative validation of the model by experts

- ✓ Appendices A through D contain the interview guides that we have used to elicit expert opinion in the process of developing and quantifying BBNs and the system dynamics model. Each interview guide was designed and used for different purposes that are explained accordingly at the beginning of the forms. These guides were used at different stages of model development and quantification.

Modeling baseline risk of specific adverse events:

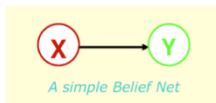
Pressure Ulcer – Model Validation

Interview Guide for First Round

Bayesian Belief Networks (BBNs): A brief introduction

Bayesian Belief Networks are a specific form of influence diagrams. BBNs are graphical models of causal relations among a set of variables (factors), where variables are represented as nodes of the graph and the interaction between the variables (causality) as arcs (directed edges) between the nodes. Any pair of unconnected nodes of such a graph indicates independence between the variables represented by nodes. Hence, BBNs, or probabilistic networks in general, capture a set of dependence and independence properties associated with the variables represented by nodes in the network. To specify the strengths of these dependence relations, we use conditional probabilities.

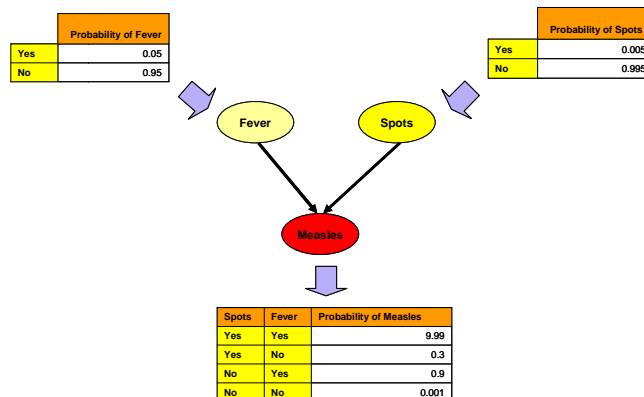
In short, BBN is a directed acyclic graph that represents causality relationships between variables and consists of:



A set of variables

A set of directed arcs linking pairs of nodes; an arc from X to Y means that X has a direct influence on Y (we call X the parent node and Y the child node)

Each node has a conditional probability table that quantifies the effects of the parents on the child node



Background

What is your position or role in (Your organization)? Please describe your daily work and/or responsibilities in your current role.

How long have you worked in the (Your unit)?

What is your professional/educational background?

Model 1-Pressure Ulcer

As you know, every hospitalized patient is vulnerable to a certain level of risk experiencing pressure ulcer. Factors that influence this risk can be categorized into:

Patient level factors (relating to patient's conditions)

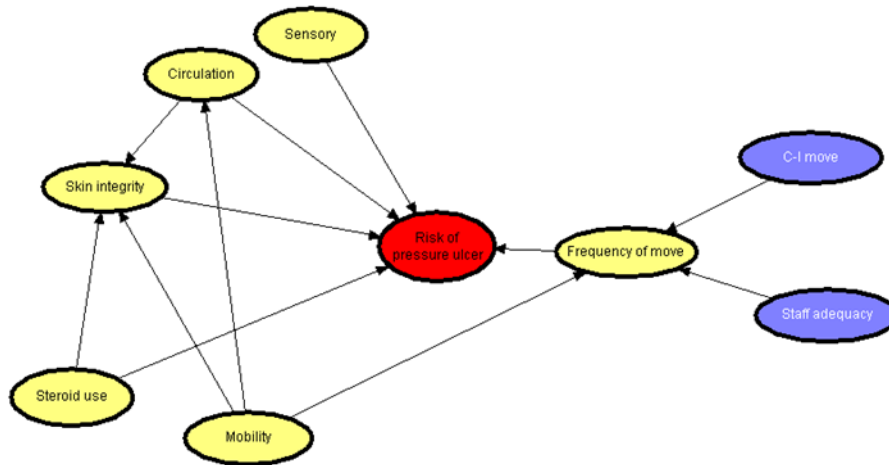
Physician-Patient level factors (relating to the treatment of the patient)

From your perspective what are the most important factors that influence the risk of experiencing pressure ulcer, while a patient is in the hospital? Based on the brief introduction provided on influence diagrams could you please sketch a diagram that shows these important factors and how they impact the risk of experiencing pressure ulcer?

Response to Base Model

Take a look at this diagram. Based on the influence diagram you provided, let's fill in parts that you mentioned, but that are missing from this model. Also, I see a few events in this diagram that you didn't mention.

[Interviewer will iteratively work with the interviewee/subject to incorporate or exclude specific variables from the base model]



Attributes:

Completeness: From your perspective, to what extent does this model capture all important and relevant phenomena for the particular problem that we are studying? On a scale from 0 to 100, 0 would correspond to a model that does not include some important and relevant details, whereas 100 would correspond to a model that includes all details that you consider important. What number would you assign?

Accuracy: From your perspective, how accurately or realistically does the model depict important factors that influence risk of experiencing pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that is unrealistic, over-idealized or inaccurate, whereas 100 would correspond to a model that is realistic and accurate. What number would you assign?

Ease of understanding: From your perspective, how easy is it to understand the overall logic of the model. On a scale from 0 to 100, 0 would correspond to a model that is difficult to follow, even with extensive explanation, and a 100 would correspond to a model that is readily understandable. What number would you assign?

Perceived predictive validity: From your perspective, if you were to use this model, how well could you predict the risk of pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that does not help at all with predicting effects, and a 100 would correspond to a model that predicts the effects very well. What number would you assign?

Final Comments

Do you have any other comments that you want to make? Feel free to elaborate on anything that we have already discussed.

Appendix B. Model Validation Interview Guide, line Infection BBN

Purpose:

- A. Introducing experts to the concept of Bayesian belief networks, the structure and the concept of conditional probabilities
- B. Introducing the preliminary BBN model for risk of line infection
- C. Eliciting expert's opinion about the factors and relations in the model, addition or deletion of the factors if necessary based on experts judgment
- D. Eliciting expert's quantitative assessment on some of the parameters of the model
- E. Qualitative validation of the model by experts

Modeling baseline risk of specific adverse events:

Line Infection – Model Validation

Interview Guide for First Round

Bayesian Belief Networks (BBNs): A brief introduction

Bayesian Belief Networks are a specific form of influence diagrams. BBNs are graphical models of causal relations among a set of variables (factors), where variables are represented as *nodes* of the graph and the interaction between the variables (causality) as *arcs (directed edges)* between the nodes. Any pair of unconnected nodes of such a graph indicates independence between the variables represented by nodes. Hence, BBNs, or probabilistic networks in general, capture a set of dependence and independence properties associated with the variables represented by nodes in the network. To specify the strengths of these dependence relations, we use *conditional probabilities*.

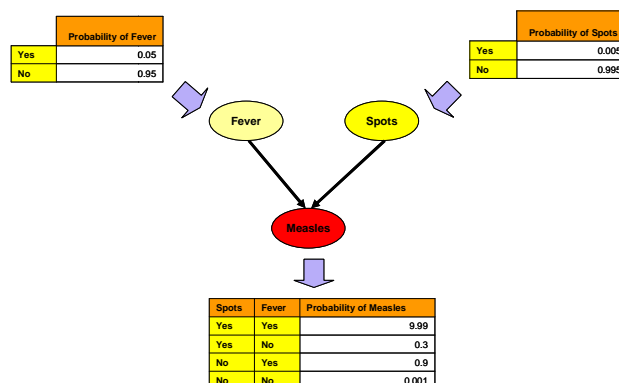
In short, BBN is a directed acyclic graph that represents causality relationships between variables and consists of:



A set of variables

A set of directed arcs linking pairs of nodes; an arc from X to Y means that X has a direct influence on Y (we call X the parent node and Y the child node)

Each node has a conditional probability table that quantifies the effects of the parents on the child node



Background

What is your position or role in (Your organization)? Please describe your daily work and/or responsibilities in your current role.

How long have you worked in the (Your unit)?

What is your professional/educational background?

Model 2-line infection

As you know, every hospitalized patient is vulnerable to a certain level of risk experiencing line infection. Factors that influence this risk can be categorized into:

Patient level factors (relating to patient's conditions)

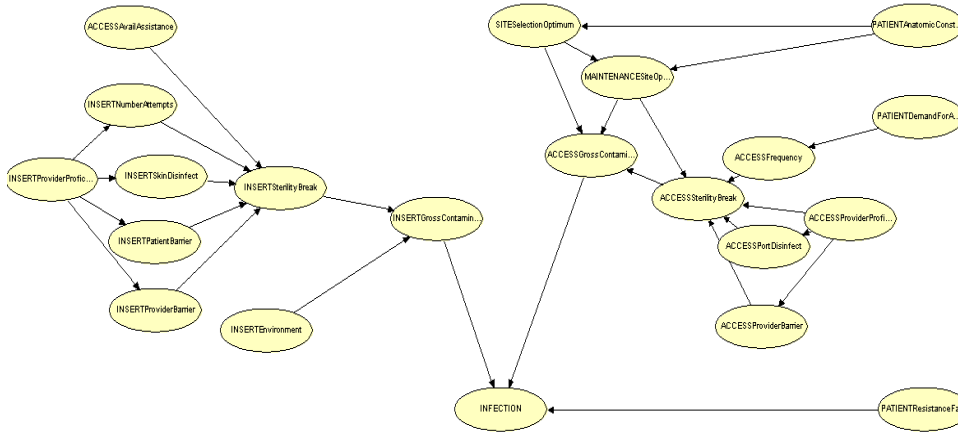
Physician-Patient level factors (relating to the treatment of the patient)

From your perspective what are the most important factors that influence the risk of experiencing line infection, while a patient is in the hospital? Based on the brief introduction provided on influence diagrams could you please sketch a diagram that shows these important factors and how they impact the risk of experiencing line infection?

Response to Base Model

Take a look at this diagram. Based on the influence diagram you provided, let's fill in parts that you mentioned, but that are missing from this model. Also, I see a few events in this diagram that you didn't mention.

[Interviewer will iteratively work with the interviewee/subject to incorporate or exclude specific variables from the base model]



Attributes:

Completeness: From your perspective, to what extent does this model capture all important and relevant phenomena for the particular problem that we are studying? On a scale from 0 to 100, 0 would correspond to a model that does not include some important and relevant details, whereas 100 would correspond to a model that includes all details that you consider important. What number would you assign?

Accuracy: From your perspective, how accurately or realistically does the model depict important factors that influence risk of experiencing pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that is unrealistic, over-idealized or inaccurate, whereas 100 would correspond to a model that is realistic and accurate. What number would you assign?

Ease of understanding: From your perspective, how easy is it to understand the overall logic of the model. On a scale from 0 to 100, 0 would correspond to a model that is difficult to follow, even with extensive explanation, and a 100 would correspond to a model that is readily understandable. What number would you assign?

Perceived predictive validity: From your perspective, if you were to use this model, how well could you predict the risk of pressure ulcer? On a scale from 0 to 100, 0 would correspond to a model that does not help at all with predicting effects, and a 100 would correspond to a model that predicts the effects very well. What number would you assign?

Final Comments

Do you have any other comments that you want to make? Feel free to elaborate on anything that we have already discussed.

*Appendix C. Model Validation Interview Guide, System Dynamics Model;
Round 1&2*

Purpose:

- A. Introducing experts to the concept of dynamics modeling; the structure and the concept of building blocks of system dynamics
- B. Introducing the hypothesis and the preliminary SD model for capturing organizational level contributors to the risk of adverse events
- C. Eliciting expert's opinion about the factors and relations in the model, addition or deletion of the factors if necessary based on experts judgment
- D. Eliciting expert's assessment on the shape and forms of the soft factors in the model (i.e. pressure functions)
- E. Qualitative validation of the model by experts

Modeling dynamic aspects of adverse events risk:

Model Validation

Interview Guide-Round 1&2

Background

What is your position or role in (Your organization)? Please describe your daily work and/or responsibilities in your current role.

How long have you worked in the (Your unit)?

What is your professional/educational background?

System Dynamics-A brief introduction

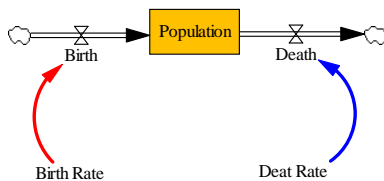
SD is an approach to modeling systems and how they change overtime. It is a simulation based, differential equation modeling tool that is used when:

Formal model is complex

Analytical solution is impossible or very difficult to obtain

It has been used in variety of problems such as corporate strategy, dynamics of diabetes, cold war arm race, HIV combat with human immune system .

The building blocks of a SD model:



Stocks: accumulation of a measureable entity

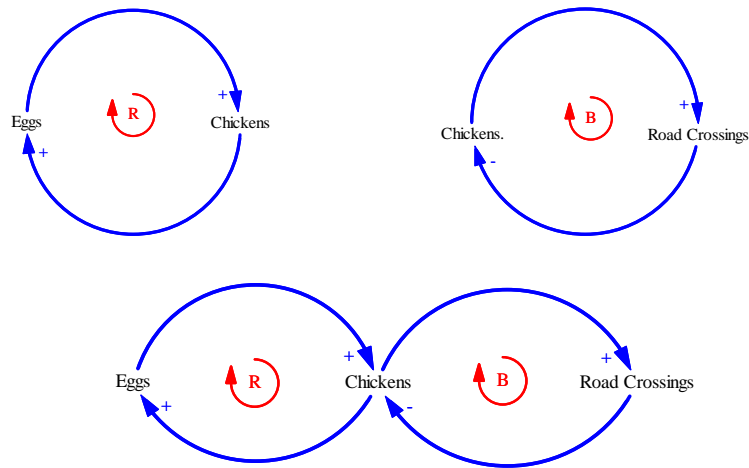
People, parts, money or intangibles such as happiness (Ford,99)

Characterize the state of the system

Generate information for decision making

Flows: Physical or conceptual entities that enter or exit system

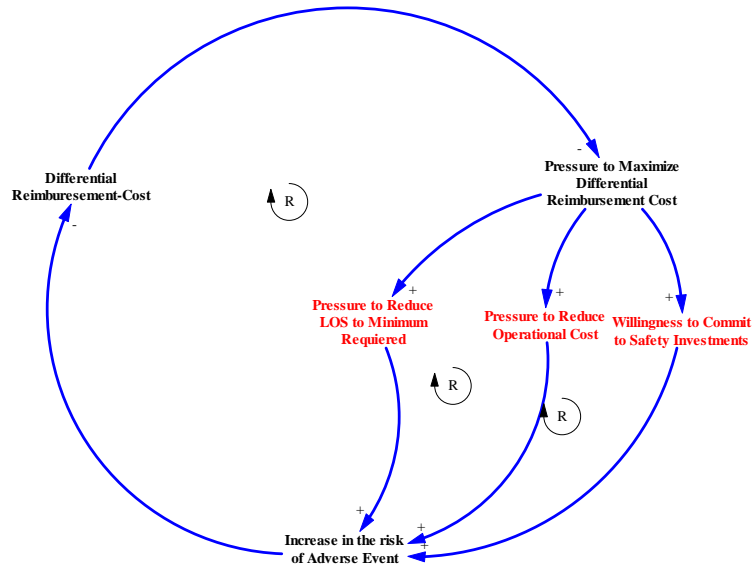
Feedback Loops



Model 3-SD

Response to Integrated Model –System Dynamics Model

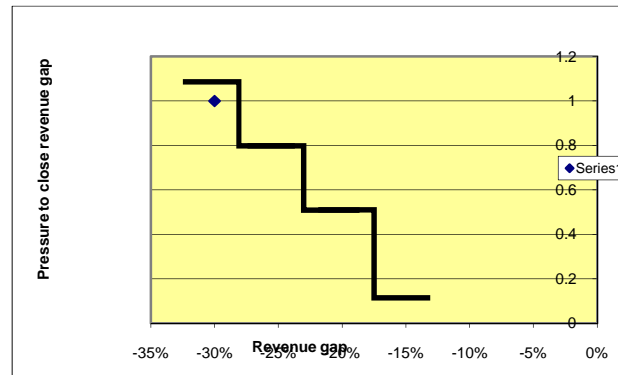
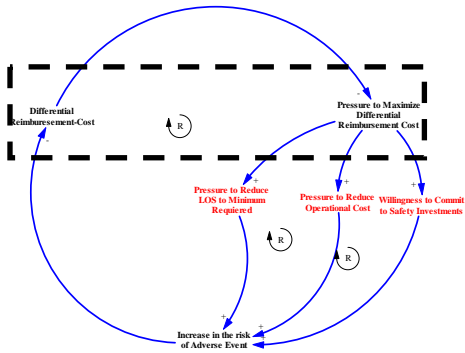
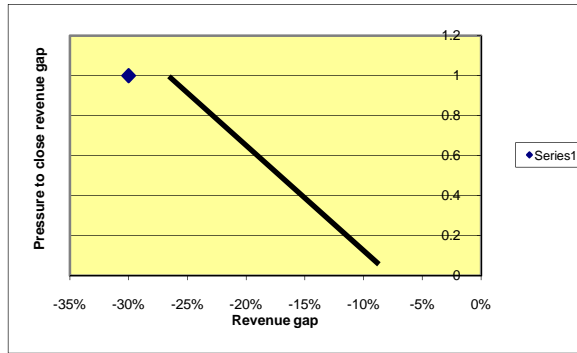
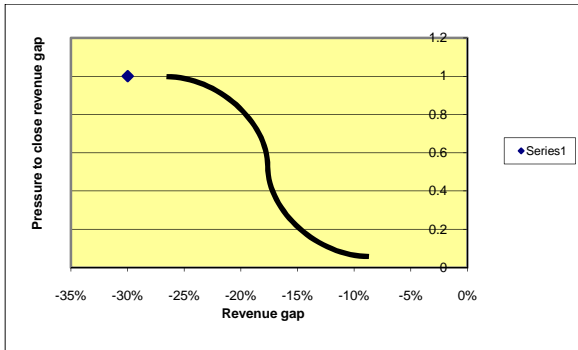
Please take a look at this diagram. This model captures the overall responses/decisions to revenue gap in US hospitals and their effect on risk of experiencing specific adverse events. It also captures the feedback effect of this risk on the same revenue gap that triggered this process. I will ask you to grade the model along several dimensions.

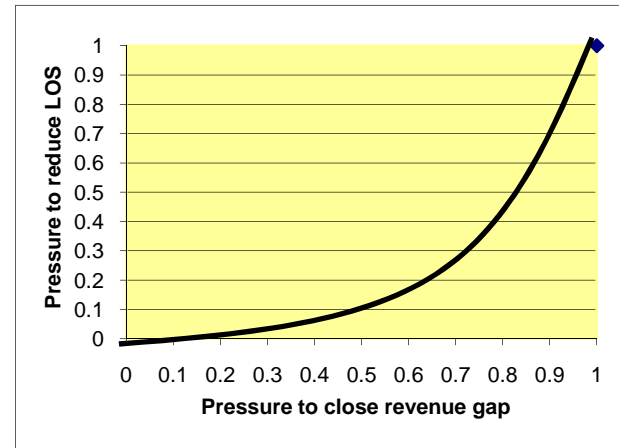
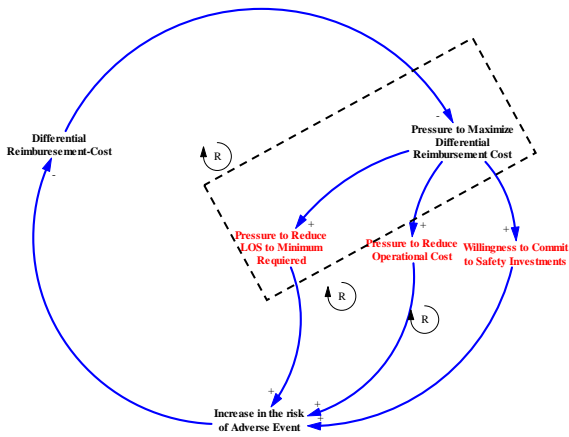
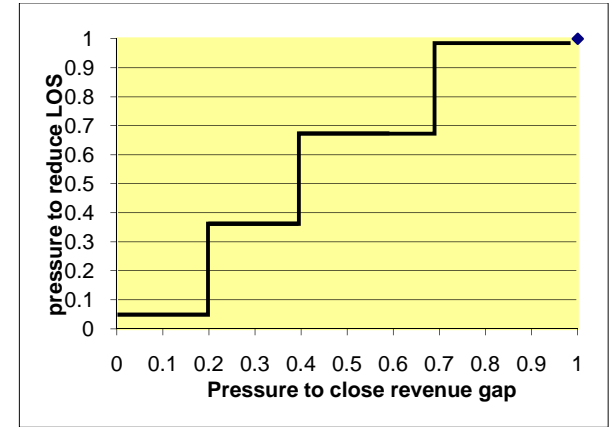
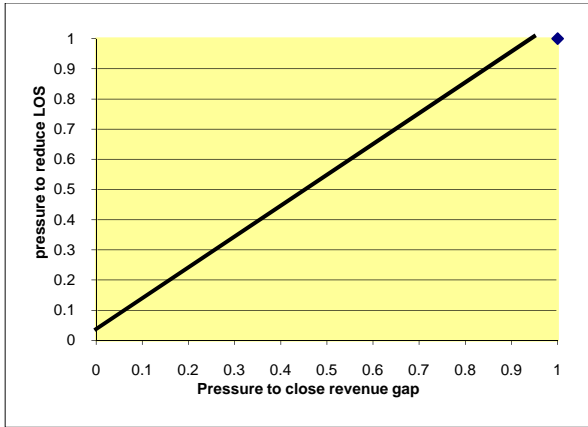


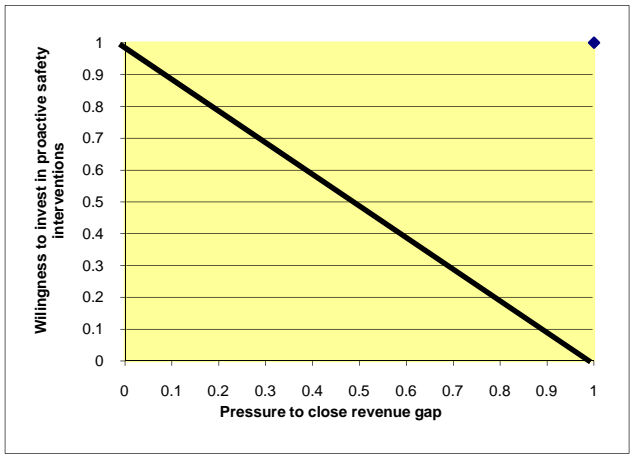
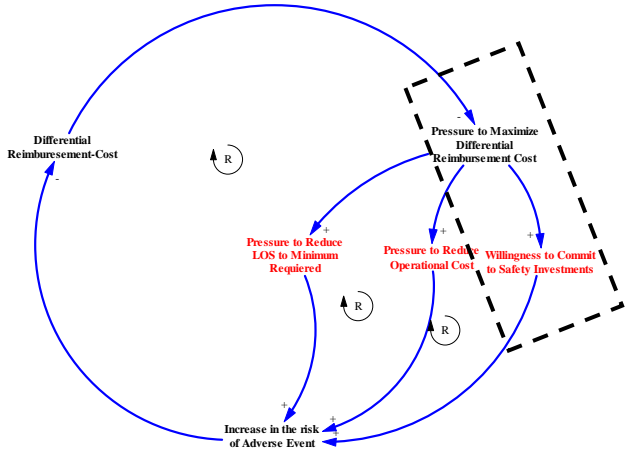
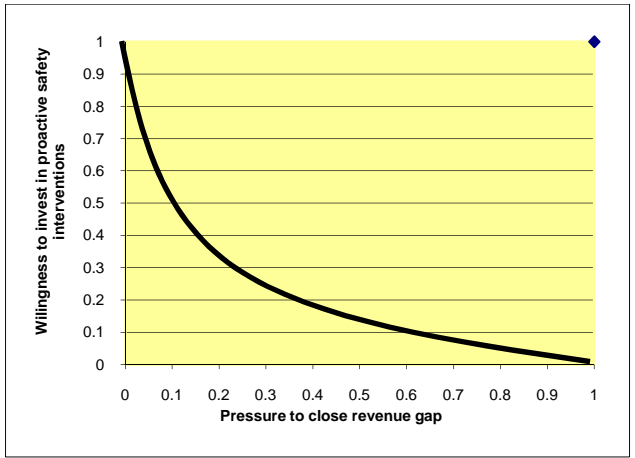
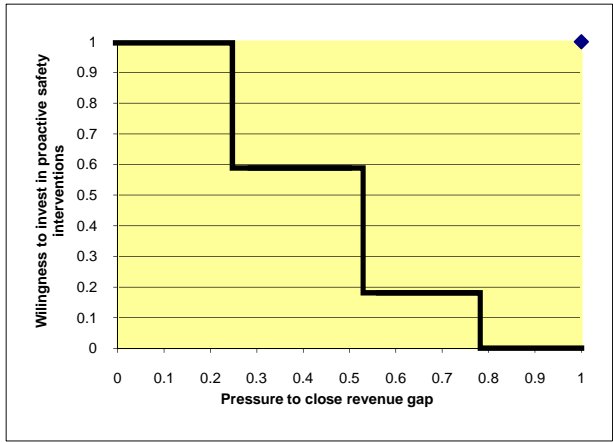
General questions:

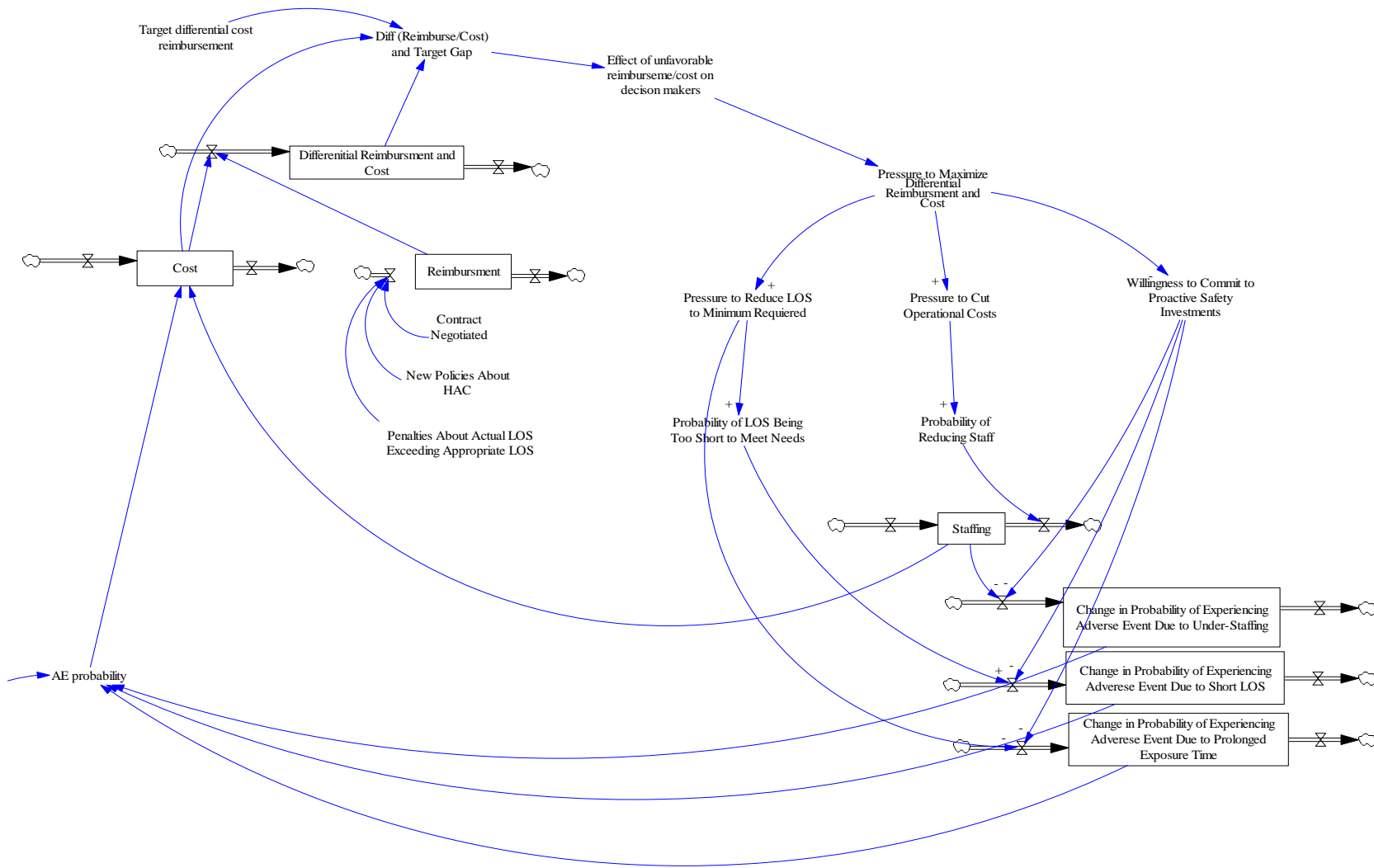
Ordering of these policies

Target revenue









Attributes:

Completeness: From your perspective, to what extent does this model capture all important and relevant phenomena for the particular problem that we are studying? On a scale from 0 to 100, 0 would correspond to a model that does not include some important and relevant details, whereas 100 would correspond to a model that includes all details that you consider important. What number would you assign?

Accuracy: From your perspective, how accurately or realistically does the model depict important feedback effects, and causal chains that influence risk of experiencing adverse events? On a scale from 0 to 100, 0 would correspond to a model that is unrealistic, over-idealized or inaccurate, whereas 100 would correspond to a model that is realistic and accurate. What number would you assign?

Ease of understanding: From your perspective, how easy is it to understand the overall logic of the model. On a scale from 0 to 100, 0 would correspond to a model that is difficult to follow, even with extensive explanation, and a 100 would correspond to a model that is readily understandable. What number would you assign?

Perceived predictive validity: From your perspective, if you were to use this model, how well could you predict the change in the risk of specific adverse events as a function of the organizational factors/decisions that influence risk of AEs? On a scale from 0 to 100, 0 would correspond to a model that does not help at all with predicting effects, and a 100 would correspond to a model that predicts the effects very well. What number would you assign?

Final Comments

Do you have any other comments that you want to make? Feel free to elaborate on anything that we have already discussed.

*Appendix D. Model Validation Interview Guide, System Dynamics Model;
Round 3*

Purpose:

- A. Eliciting expert's quantitative assessment on some of the parameters of the system dynamics model

Modeling dynamic aspects of adverse events risk:

Model Validation

Interview Guide-Round 3

You have previously expressed your opinion about the way the revenue gap creates the pressure to close this gap throughout the organization, and how this pressure manifests itself in the forms of “Pressure to Optimize LOS”, “Pressure to Reduce Operational Costs” and” Willingness to Invest in Proactive Safety Investments”. Considering the model above, please answer the following questions.

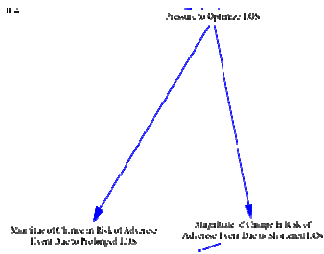
As was discussed in other rounds of interview, the pressure to optimize LOS, may affect risk of adverse events in two ways, first, it may increase the probability of experiencing an adverse events, because some patient’s LOS may be too short to meet all his/her needs. Second, it may reduce the probability of some adverse events because it simply reduces the exposure and if the pressure to optimize LOS is too small, some patients may stay in the hospital longer than they really need and be exposed to certain adverse events.

To your experience at what level (range) of pressure to optimize LOS, we might start to see the effect of this pressure on risk, because the pressure is too high that some patients may be discharged a bit prematurely?

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1
---------------	-------	---------	---------	---------	---------	---------	---------	---------	---------	-------	-----------------

To your experience at what level (range) of pressure to optimize LOS, we might start to see the effect of this pressure on risk, because the pressure is too low that some patients may stay longer than they really need to which may increase their risk of being exposed to and experiencing certain adverse events?

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1
---------------	-------	---------	---------	---------	---------	---------	---------	---------	---------	-------	-----------------

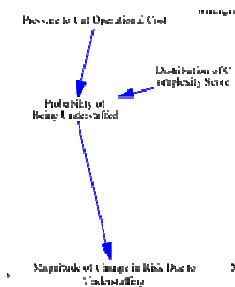


2- The pressure to reduce operational costs, may affect risk of adverse events due to understaffing.

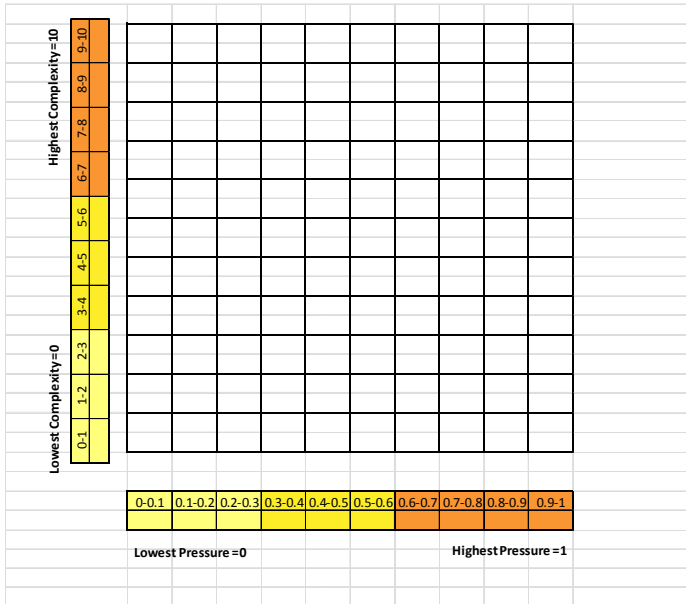
To your experience at what level (range) of pressure to reduce operational costs are great enough such that some patients may experience an adverse event because the unit is not sufficiently staffed (either due to lower numbers of staff, or lower quality of staff?)

*You may use an arrow to indicate a precise point or circle one of the ranges above to indicate a broader range estimate.

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1
---------------	-------	---------	---------	---------	---------	---------	---------	---------	---------	-------	-----------------



The idea is that the impact of lower staffing numbers and/or less experienced staff may depend on the complexity of the case mix. With this assumption in mind, suppose we take the average of the complexity scores across all units in the hospital and all inpatient days. Please indicate in the table below, where the pressure to reduce operational costs/staffing begins to influence the probability of an adverse event as a function of the complexity of the patient population.



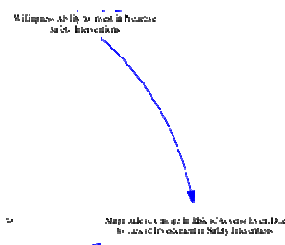
Assume that the pressure to reduce the revenue gap, will affect the level of willingness/ability to invest in proactive safety investments. The worse the financial situation gets, the less investments are made in safety programs. Assume that the more we spend on safety the less the chances of experiencing adverse events will be. If this willingness to invest in safety is a scale between 0-1 (0 meaning no ability/willingness to invest in safety, and 1 meaning highest level of ability/willingness to invest in proactive safety interventions);

Based on your experience at what level (range) of this willingness do we start to see changes in the risk of adverse event?

*You may use an arrow to indicate a precise point or circle one of the ranges above to indicate a broader range estimate.

No Pressure=0	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	Max Pressure =1
---------------	-------	---------	---------	---------	---------	---------	---------	---------	---------	-------	-----------------

Based on your experience, what is the magnitude of change in the risk of adverse events when there is an increase or decrease in investment in elective/proactive safety programs? Use the table below to indicate the relationship between changes in investments and magnitude of effect on risk of adverse events.



References

1. Allman, R., Goode, P., Patrick, M., Burst, N., Bartolucci, A., 1995, Pressure Ulcer Risk Factors Among Hospitalized Patients With Activity Limitation, JAMA, Vol. 273, No. 11, PP. 865- 870
2. Ammann, RA., Bodmer, N., Hist, A., Niggli, FK., Nadal, D., Simon, A., Ozsahin, H., Kontny, U., Kuhne, T., Popovic, MB., Luthy, AR., Aebi, C., 2010, Predicting adverse events in children with fever and chemotherapy-induced neutropenia: the prospective multicenter SPOG 2003 FN study, Journal of clinical oncology, Vol. 28, No. 12, PP. 2008-2014
3. Anders, J., Heinemann, A., Leffmann, C., Leutenegger, M., Profener, F., Renteln-Kruse, W., 2010, doi:10.3238/arztebl.2010.0371
4. Apkon, M, Leonard, J, Probst, L, DeLizio, L, Vitale, R, Design of a safer approach to intravenous drug infusions: failure mode and effect analysis, 2004, Quality and safety in healthcare, 13:265-271
5. Ayello, E., and Barden, B., 2002, How and Why to do Pressure Ulcer Risk Assessment, Advances in Skin & Wound Care, Vol. 15, No. 3, PP. 125- 133
6. Ayyub, B., 2001, Elicitation of expert opinion for uncertainty and risk, CRC press
7. Baldi, L., Ferrando, A., Foltran, f., Ciccone, G., Gregori, D., 2010, Studying Factors Related to Pressure Ulcers Prevention: A Marginal Scale Model for Modelling Heterogeneity Among Hospitals, Journal of Evaluation in Clinical Practice, doi:10.1111/j.1365-2753.2009.01257.x

8. Barden scale for Preventing Pressure Sore Risk, www.bradenscale.com
9. Barlas, Y., 1994, Model Validation in System Dynamics, Proceedings of the 1994 International System Dynamics Conference, System Dynamics Society, Stirling, Scotland, PP.1-10
10. Barlas, Y., 1996, Forma Aspects of Model Validity and Validation in System Dynamics, System Dynamics Review, Vol. 12, No. 3, PP. 183-210
11. Battista ORsi, G., Di Stefano, L., Noah, N., 2002, Hospital-Acquired, Laboratory-Confirmed Bloodstream Infection: Increased Hospital Stay and Direct Costs, Infection Control and Hospital Epidemiology, Vol. 23, No. 4, PP. 190-197
12. Baumgarten, M., Margolis, D., Berlin, J., Storm, B., Garino, J, Kagan, S., ..., 2003, Risk Factors for Pressure Ulcers Among Elderly Hip Fracture Patients, Wound Repair and Regeneration, Vol. 11, No. 2, PP. 96-103
13. Beckrich, K., Aronovitch, S., 1999, Hospital-Acquired Pressure Ulcers: A Comparison of Costs in Medical vs. Surgical Patients, Nursing Economics, Vol. 17, No. 5, PP.263-271
14. Bedford, T., Cooke, R., 2001, Probabilistic Risk Analysis; Foundations and Methods, Cambridge
15. Bennett, G., Dealey, C., Posnett, J., 2004, The Cost of Pressure Ulcers in the UK, Age and Aging, Vol. 33, no. 3, PP. 230-235
16. Bergstrom, N., Braden, B., Laguzza, A., Holman, V., 1987, The Braden scale for Predicting Pressure sore Risk, Nurs Res, Vol. 36, PP. 205-210
17. Bergstrom, N., Kemp, M., Ruby, E., 1996, Multi-site Study of Incidence of pressure Ulcers and the Relationship Between Risk Level, Demographic

Characteristics , Diagnoses, and Prescription of Preventive Interventions, Journal of American Geriatrics Society, Vol. 44, PP. 22-30

18. Bogner, M, A system approach to medical error. In Safety in medicine, 2001, 83-100, Pergamon
19. Bonnabry, P, Cingaria, L, Sadeghipoue, F, Ing, H, Fonzo-Christe, C, Pfister, R, Use of a systematic risk analysis method to improve safety in the production of pediatric parenteral nutrition solution, 2005, Quality and safety in healthcare, 14:93-98
20. Borlawsky, T., 2004, pressure Ulcer Risk assessment Model, Columbia University, Department of Biomedical Informatics
21. Brennan, T.A., Leape L.L., Larid N., Herbert L., Localis A.R., Lawthers A.G., Newhouse J.P., Weiler P.C., and Hiatt H.H., Incidence of adverse events and negligence in hospitalized patients: Results of the Harvard medical practice study, 1991, New England Journal of Medicine, 324, 370-376
22. Brown, S., Development and validation of methodology for fix effectiveness projection during product development, 2009, PhD thesis, University of Maryland, College Park
23. Burgemier, J, Failure mode and effect analysis: an application in reducing risk in blood transfusion, 2002, Joint commission journal on quality improvement, 28: 331-339
24. Calvillo-King, L., Xuan, L., Zhang, S., Tuhrim, S., Halm, EA., 2010, Predicting risk of perioperative death and stroke after carotid endarterectomy in

asymptomatic patients; derivation and validation of a clinical risk score, *Stroke*,
Vol.41, No. 12, PP. 2786-2794

25. Carson, J., 2002, Model Verification and Validation, Proceedings of the 2002 Winter Simulation Conference, PP. 52-58
26. Carthey, J, De Leval, M, Reason, J, Leggatt, A , Wright, D, 2000, Adverse events in cardiac surgery: the role played by human and organizational factors. In *Safety in medicine*, PP. 117-138, Pergamon
27. Chadwick, L., Fallon, E., 2011, Human reliability assessment of a critical nursing task in a radiotherapy treatment process, *Applied ergonomics*,
[doi:10.1016/j.apergo.2011.03.011](https://doi.org/10.1016/j.apergo.2011.03.011)
28. Comden, S, Marx, D, Murphy-Carely, M, Hale, M, Using probabilistic risk assessment to model medication system failure in long-term care facilities, 2005, *Advances in patient safety*, 2:395-408
29. Cook, D.L., 2004, The Dynamics and Control of Operational Risk, Ph.D. Thesis, University of Calgary, Calgary, Alberta
30. Cooke, R., 1991, *Experts in Uncertainty*, Oxford University Press, New York/Oxford
31. Cox, A., Dolan, L., MacEwan, CJ., 2008, Human reliability analysis: a new method to quantify errors in cataract surgery, *Eye*, Vol. 22, PP. 394-397,
[doi:10.1038/sj.eye.6702648](https://doi.org/10.1038/sj.eye.6702648)
32. Coyle, G., 1977, *Management System Dynamics*, Chichester, John Wiley and Sons

33. Coyle, G., Exelby, D., 2000, The Validation of Commercial System Dynamics Models, *System Dynamics Review*, Vol. 16, No. 3, PP. 225-244
34. Cromheeche, Koornneef, F, Van Gaalen, G, de Mol, B, Controlling the risks of mechanical heart valve failure using product life cycle based safety management,. In *Safety in medicine*, 2000, 175-192, Pergamon
35. Defloor, T., Grypdonck, M., 2004, Pressure Ulcers: Validation of Two Risk Assessment Scales, *Journal of Clinical Nursing*, Vol. 14, PP. 373-382
36. DeRosier, J, Stalhandske, E, Bagain, J, Nudell, T, Using healthcare failure mode and effect analysis: the VA national center for patient safety's prospective risk analysis system, 2002, *Joint commission journal on quality improvement*, 28:248-267
37. Dierks, M, Dulac, N, Levenson, N, System dynamics approach to model risk in complex healthcare settings: time constraints, production pressures and compliance with safety controls 2007,
38. Diez, J., 1993, Parameter Adjustment in Bayes Netwroks, The Generalized Nosi-OR Gate, *Proceedings of the Ninth annual Conference on Uncertainty in Artificial Intelligence*, Washington DC., PP.99-105
39. Digiovine, B., Chenoweth, C., Watts, C., Higgins, M., 1999, The Attributable Mortality and Costs of Primary Nosocomial Bloodstream Infections in the Intensive Care Unit, *Am J Respir Crit Care Med*, Vol. 160, PP. 976-981
40. Droguett, E., 1999, Methodology for the treatment of model uncertainty [PhD thesis], College park, University of Maryland

41. Droguett, E., and Mosleh, A., 2008, Bayesian methodology for model uncertainty using model performance data, *Risk analysis*, Vol. 28, Issue 5, Pages 1457-1476
42. Drudzel, M., van der Gaag, L.C., 2000, Building Probabilistic Networks,: Where Do the Numbers Come From?, Guest Editor's Introduction,, *IEEE transactions on Knowledge and Data Engineering*, Vol. 12, No.4, PP. 481-486
43. Druzdzel, M. J. and L. C. van der Gaag, 2000, Building Probabilistic Networks: Where Do the Numbers Come From?, *IEEE transactions on knowledge and data engineering*, Vol. 12, No. 4, PP. 481-486
44. Druzdzel, M. J. and L. C. van der Gaag: 1995, Elicitation of Probabilities for Belief Networks: Combining Qualitative and Quantitative Information'. In: P. Besnard and S. Hanks (eds.): *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. San Francisco: Morgan Kaufmann, PP. 141-148.
45. Fahlbruch, B, Wilpert, B, Vincent, C, Approaches to safety. In *Safety in medicine*, 2000, PP.9-29, Pergamon
46. Ford, A., *Modeling the environment: an introduction to system dynamics modeling of environmental systems*, 1999, Washington DC., Island press
47. Forrester, J., *Collected papers*, 1975, Wright-Allen Press
48. Forrester, J., Senge, P., 1980, Tests for Building Confidence in System Dynamics Models, *TIMS studies in the Management Sciences*, Vol. 14, PP. 209-228
49. Fortinsky, R., Iannuzzi-Suicich, M., Baker, D., Gottschalk, M., King, M., Brown, C., Tinetti, M., 2004, *Fall-Risk Assessment and Management in Clinical Practice*:

Views From Healthcare Providers, Journal of American Geriatrics Society, Vol. 52, Issue 9, PP. 1522-1526

50. Garcia-Garmendia, J., Ortiz-Leyba, C., Garnacho-Montero, J., Jimenez, F., Monterrubio, J., Gili-Miner, M., 1999, Mortality and the Increase in Length of Stay Attributable to the Acquisition of Acinetobacter in Critically Ill Patients, Critical Care Medicine, Vol. 27, No. 9, PP. 1794-1799
51. Gauba, V., Tsangaris, P., Tossounis, C., Mitra, A., McLean, C., Saleh, G., 2008, human reliability analysis of cataract surgery, Archives of ophthalmology, Vol. 129, No., 6, PP.173-177
52. Gosnell, DJ. 1973, An Assessment Tool to Identify Pressure Sores. Nurs Res, Vol.22, No.1, PP.55-59
53. Grundy, S., Pasternak, R., Greenland, P., Smith, S., Fuster, V., 1999, Assessment of Cardiovascular Risk by Use of Multiple Risk Factor assessment Equations, Circulation, Vol. 100, PP. 1482-1492
54. Gunningberg, I., 2005, Are Patients With or at Risk of Pressure Ulcers allocated Appropriate Prevention Measures?, International Journal of Nursing Practice, Vol. 11, PP. 58-67
55. Hannon, B., Ruth, M., 2001, Dynamic Modeling, Springer, Second Edition
56. Heckerman, D., Breese, J., 1996, Causal Independence for probabilistic Assessment and Inference Using Bayesian Networks, IEEE Transactions on Systems, man and Cybernetics, Vol. 26, PP. 826-831

57. Hnrion, M., 1989, Some Practical Issues in Constructing Belief Networks, Uncertainty in Artificial Intelligence 3, eds., L.N. Kanal, T.S. Levitt, and J.F. Lemmer, Elsevier Science Publishers, North Holland, PP. 161-173
58. Hodge, R., Evans, M., Marshall, J., and Walls, L., Eliciting engineering knowledge about reliability during design: lessons learnt from implementation, 2001, Quality and reliability engineering international, 17:169-179
59. Hollnagle, E., CREAM: cognitive reliability error analysis method, 1998, Esilver London
60. Hollnagle, E., Human reliability analysis: context and control, 1993, Academic press, London
61. Hyman, W, Why don't healthcare professionals do what they know they should?, 2005, Journal of clinical engineering, 214-218
62. Iglesia, C., Nixon, J., Cranny, G., Nelson, A., Hawkins, K., Philips, A., ..., 2006, Pressure Relieving Support Surfaces (PRESSURE) Trial: Cost Effectiveness Analysis, BMJ, doi: 10.1136/bmj.38850.711435.7C
63. Inoue, k., Koizumi, A., 2004, Application of human reliability analysis to nursing errors in hospitals, Risk analysis, Vol. 24, No. 6, PP. 1459-1473
64. Jiricka, M., Ryan, P., Carvalho, M., Bukvich, J., 1995, Pressure Ulcer Risk Factors in an ICU Population, American Journal of Critical Care, Vol. 4, No. 5, PP. 361-367
65. Joice, P., Hanna, G.B, Cushcieri, A., 1998, Errors enacted during endoscopic surgery-a human reliability analysis, Applied ergonomics, Vol. 29, No. 6, PP. 409-414

66. Keller, B., 2006, Risks and Risk-Analysis for the Development of Pressure Ulcers in Surgical Patients, Utrecht University, The Netherlands
67. Kim, JS., Holtom, P., Vigen, C., 2011, Reduction of Catheter-Related Bloodstream Infections Through the Use of Central Venous Line Bundle: Epidemiologic and Economic Consequences, American Journal of Infection Control, E.Pub ahead of print
68. Kjaerulff, U., Madsen, A., 2008, Bayesian Networks and Influence Diagrams, Springer
69. Kohn, L, Corrigan, J, Donaldson, M, To err is human, 2000, Institute of medicine
70. Kuhn, B., Coulter, S., 1992, balancing the Pressure Ulcer Cost and Quality Equation, Nursing Economics, Vol. 10, no. 5, PP. 353-359
71. Leape, L, Error in medicine, JAMA ,1994, 272:1851-1857
72. Leape, L.L., A system analysis approach to medical error, 1997, Journal of evaluation in clinical practice, 3:213-222
73. Leape, L.L., Brennan T.A., Larid N., Lawthers A.G., Localis A.R., Barnes B.A., Herbert L., Newhouse J.P., Weiler P.C., and Hiatt H.H., The nature of adverse events in hospitalized patients: Results of the Harvard medical practice study 2, 1991, New England Journal of Medicine, 324, 377-384.
74. Lemmer, J., Gossink, D., 2004, Recursive Noisy OR—A Rule for Estimating Complex Probabilistic Interactions, IEEE Transactions on Systems, Man and Cybernetics, Vol. 34, No. 6, PP. 2252- 2261

75. Lindgren, M., Unosson, Krantz, A., Ek, A., 2004, Pressure Ulcer risk Factors in Patients Undergoing surgery, *Journal of Advanced Learning*, Vol. 50, No. 6, PP. 605-612
76. Lindgren, M., Unosson, M., Fredrikson, M., Ek, A., 2004, Immobility- A Major Risk Factor for Development of Pressure Ulcers Among Adult Hospitalized Patients: A Prospective Study, *Scand j Caring Sci*, Vol. 18, PP. 57-64
77. Lyder, C., 2003, pressure Ulcer Prevention and Management, *JAMA*, Vol. 289, No. 2, PP. 223-226
78. Lyder, C., preston, J., Grady, J., Scinto, J., Allman, R., Bergstorm, N., Rodeheaver, G., 2001, Quality of Care for Hospitalized Medicare Patients at Risk for Pressure Ulcers, *Arch Internal Medicine*, Vol. 161, PP. 1549- 1554
79. Lyons, M, Adams, S, Woloshynowych, M, Vincent, C, 2004, Human reliability analysis in healthcare: a review of techniques, *International journal of risk and safety in medicine*, 16, PP.223-237
80. Maglogiannis, I, Zafiropoulos, E, Modeling risk in distributed healthcare information systems, 2006, *Proceeding of the 28th IEEE , EMBS annual international conference*
81. Mahieu, L., Muynck, A., Leven, M., Dooy, J., Goossens, H., Van Reempts, P., 2001, Risk Factors for Central Catheter-Associated Bloodstream Infections Among Patients in a Neonatal Intensive Care Unit, *Journal of Hospital Infection*, Vol. 48, PP.108-116

82. Malik, R., White, P.S., Macewen, C.J., 2003, Using human reliability analysis to detect surgical error in endoscopic DCR surgery, *Clinical otolaryngology and allied sciences*, Vol. 28, No. 5, PP.456-460
83. Marcot, B., Steventon, D., Sutherland, G., McCann, R., 2006, Guidelines for Developing and Updating Bayesian Belief Networks Applied to Ecological Modeling and Conservation, *Can. J. For. Res.*, Vol. 36, PP. 3063-3074
84. Margolis, D., Knauss, J., Bilker, W., Baumgarten, M., 2003, Medical Conditions as Risk Factors for Pressure Ulcers in an Outpatient Setting, *Age and Aging*, Vol. 32, PP. 259-264
85. McCabe, R., 1998, *Why Airlines Succeed or Fail: A System Dynamics Synthesis*, Ph.D. Thesis, Claremont Graduate University, Claremont, CA
86. McInnes, E., Cullum, N.A., Bell-Syer, S, Dumville, J., Jammali-Blasi, A., 2010, Support Surfaces for Pressure Ulcer Prevention (Review), *Cochran Library*, Issue 5
87. McLaughlin, C., Kaluzny, A., Total quality management in health: making it work, 1990, *Healthcare management review*, 15(3): 7-14
88. McNally, K, Page, M, Sunderland, V, Failure mode and effect analysis in improving a drug distribution system, 1997, *American journal of health system policy*, 54:171-177
89. Modarres, m., 2006, *Risk Analysis in Engineering: techniques, Tools, and Trends*, Boca Raton, FL., Taylor Francis Group
90. Mohagheh, Z, *On the Theoretical Foundations and Principles of Organizational Safety Risk Analysis*, 2007, University of Maryland, College Park

91. Mohaghegh, Z., Kazemi, R., Mosleh, A., Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization, 2008, Reliability engineering and system safety, 94 (5):1000-1018
92. Moro, M., Vigano, E., Lepri, A., 1994, Risk Factors for Central Venous Catheter-Related Infections in Surgical and Intensive Care Units, Infection Control and Hospital Epidemiology, Vol. 15, No. 4, PP.253-265
93. Mosleh, A., Apostolakis, 1986, The Assessment of probability Distributions from Experts Opinions With an Application to Seismic Fragility Curves, Risk Analysis, Vol. 6, No. 4, PP. 447-461
94. Mosleh, A., Apostolakis, G., The assessment of probability distributions from expert opinions with an application to seismic fragility curves, 1986, Risk analysis, 6: 447-461
95. Mosleh, A., Bier, V., Apostolakis, G., A critique of current practice for the use of expert opinions in probabilistic risk assessment, 1988, Reliability engineering and system safety, 20: 63-85
96. Mosleh, A., Lecture notes, Advanced methods in reliability engineering, University of Maryland, College Park
97. Mosleh A., and V.M. Bier, *On Decomposition and Aggregation Error in Estimation: Some Basic Principles and Examples*, 1992, Risk Analysis, 12: 203-214.
98. Mrdovic, I., Savic, L., Krljanac, g., Asanin, M., Perunicic, J., Lasica, R., Marinkovic, J., Kocev, J., Vasiljevic, Z., Ostojic, M., 2011, Predicting 30-day

major adverse cardiovascular events after primary percutaneous coronary intervention. The risk-PCI score, *International journal of cardiology*,
[doi:10.1016/j.ijcard.2011.05.071](https://doi.org/10.1016/j.ijcard.2011.05.071)

99. Neil, M., Fenton, N., Nielson, L., 2000, Building Large Scale Bayesian Networks, *The Knowledge Engineering Review*, 15 (3), PP. 257-284
100. Nixon, J., Cranny, G., Iglesias, C., Nelson, E., Hawkins, k., Philips, A., ..., 2006, Randomised, Controlled Trial of Alternating Pressure Mattresses Compared with Alternating Pressure Overlays for the Prevention of Pressure Ulcers: PRESSURE (Pressure Relieving Support Surfaces) Trial ,*BMJ*,[doi:10.1136/bmj.38849.478299.7c](https://doi.org/10.1136/bmj.38849.478299.7c)
101. Norton, D., McLaren, R., smith, A., 1962, An Investigation of Geriatric Nursing Problems in Hospital, London: National Corporation for the Care of Old People, PP. 238
102. Nyssen, A.S., 2001, Analysis of human errors in Anesthesia, In *Safety in medicine*, PP. 49-64, Pergamon
103. O'Hagan A., Buck, C., Daneshkhah, A., Eiser, R., Garthwaithe, P., Jenkinson, J., Oakley, J., Rakow, T., 2006, *Uncertain Judgments: Eliciting Experts' Probabilities*, John Wiley and Sons
104. Onisko, A., Druzdzel, M., Wasyluk, H., 2001, Learning Bayesian Network parameters from Small Data Sets: Application of Noisy-OR Gates, *International Journal of Approximate Reasoning*, Vol. 27, 2, PP. 165-182
105. Ostrom, L, Leahy, T, Novack, S, Summary of 1991-1992 misadministration event investigations, 1994

106. Otway, H., Winterfeldt, D., 1992, Expert Judgment in Risk Analysis and Management: process, Context, and Pitfalls, Risk Analysis, Vol. 12, No. 1, PP. 83-93
107. Ovretveit, J., Total quality management in European healthcare, 2000, International journal of healthcare quality assurance, 13(2): 74-79
108. Pandorbo-Hidalgo, P., Garcia-Fernandez, F., Lopez-Medina, I., Alvarez-Nieto, C., 2005, Risk Assessment Scales for Pressure Ulcer Prevention: A Systematic Review, Journal of Advanced Nursing, Vol. 54, No. 1, PP.94-110
109. Papanikolaou, P., Lyne, P., Anthony, D., 2007, Risk Assessment scales for Pressure Ulcers: A Methodological Review, International Journal of Nursing studies, Vol. 44, PP. 285-296
110. Pate-Cornell, E, Lakats, L, Murphy, D, Anesthesia patient risk: a quantitative approach to organizational factors and risk assessment, 1997, Risk analysis, 17: 511-523
111. Payne, N., Carpenter, J., Badger, G., Horbar, J., Rogowski, J., 2004, Marginal Increase in Cost and Excess Length of Stay Associated With Nosocomial Bloodstream Infections in Surviving Very Low Birth Weight Infants, Pediatrics, Vol. 114, No.2, PP. 348-355
112. Pearl, J., 1988, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers, San Mateo, CA
113. Pederson, C.A., Schneider, P.J., Scheckelhoff, D.J., 2002, American Journal of Health-System Pharmacy, Vol. 60, No.1, PP. 52-68

114. Perneger, T., Heliot, C., rae, A., Borst, F., Gaspoz, J., 1998, Hospital-Acquired Pressure Ulcers, Arch Internal Medicine, Vol. 158. PP. 1940-1945
115. Phipps, D., Parker, D., Meakin, G., Beatty, P., 2010, determinants of intention to deviate from clinical practice guidelines, Ergonomics, Vol. 53, No. 3, PP. 393-403
116. Rasmussen, J., Risk management in a dynamic society: a modeling problem, 1997, Safety science, 27: 183-213
117. Reason, J, Human error, 1990, Cambridge University Press
118. Reason, J, Understanding adverse events: human factors, Quality and safety in healthcare, 1995,4:80-89
119. Reason, J., Understanding adverse events: human factors, 1995, Quality and safety in healthcare, 4: 80-89
120. Reddy, M., Gill, S., & Rechon, p.,2006, preventing pressure Ulcers: A Systematic Review, Journal of the American Medical Association (JAMA), Vol. 296, No.8, PP. 974-984
121. Richet, H., Hubert, B., Nitemberg, G., Andremont, A., Buh-hoi, A., Ourbak, P., Galicier, C., et al., 1990, Prospective Multicenter Study of Vascular-Catheter-Related Complications and Risk Factors for Positive Central-Catheter Cultures in Intensive Care Unit Patients, Journal of Clinical Microbiology, Vol. 28, No. 11, PP. 2520-2525
122. Robinson, S., 1997, Simulation Model Verification and Validation: Increasing the User's Confidence, Proceedings of 1997 Winter Simulation Conference, PP. 53-59

123. Saizy-Callaert, S, Causse, R, Thebault, A, Chouaid, C, Analysis of mode of failure, their effects and criticality: improving of the hospital drug prescribing process, 2001, *Therapie*, 56:525-531
124. Sargent, R., 2001, Some Approaches and Paradigms for Verifying and Validating Simulation Models, Proceedings of the 2001 Winter Simulation Conference, ed. B.A. Peters, J.S. Smith, D. J., Medeiros, and M. W. Rohrer, PP.106-114, Piscataway, NJ
125. Seongsook, J., Ihnsook, J., Younghee, L., 2003, Validity of Pressure Ulcer Risk Assessment scales; Cubbin and Jackson, Braden, and Douglas scale, *International Journal of Nursing Studies*, Vol. 41, PP. 199-204
126. Serpa, I., Santos, V., Peres, G., Cavicchioli, M., Hermida, M., 2010, Validity of the Braden and Waterlow Subscales in Predicting Pressure Ulcer Risk in Hospitalized Patients, *Applied Nursing Research*, In Press
127. Smith, C., 2011, Integrated Scenario-Based Methodology for Project Risk Management, Ph.D. Thesis, University of Maryland, College Park
128. Smith, D., 1995, Pressure Ulcers in Nursing Homes, *Annals of Internal Medicine*, Vol. 123, No. 6, PP. 433-438
129. Smith, L., Booth, N., Douglas, D., Robertson, W., Walker, A., Durie, M., ..., 1995, A Critique of 'at Risk' pressure sore Assessment Tools, *Journal of Clinical Nursing*, Vol. 4, No. 3, PP. 153-159
130. Staender, S., Kaufmann, M., Scheidegger, D., 2001, Critical incident reporting, In *Safety in medicine*, PP. 65-82, Pergamon

131. Stamatelatos, M., Probabilistic risk assessment procedures guide for NASA managers and practitioners, 2002, NASA
132. Stanphone, N., Vincent, C., Adams, S., O'Connor, A.M., Beard, RW, Applying human factors methods to clinical risk management in obstetrics, British journal of obstetrics and gynecology, 1997, 104: 1225-1232
133. Steinwald, B., Dummit, L., 1989, Hospital case-Mix Change: Sicker patients or DRG Creep? , Health Affairs, Vol. 8, No. 2, PP. 35-47
134. Serman, J., Business Dynamics, 2000, Irwin/McGraw hill
135. Tang, Z., McCabe, B., Developing complete conditional probability tables from fractional data for BBNs, 2007, Journal of computing in civil engineering, 21 (4): 265-276
136. Taylor -Adams, S, Vincent, C, Clinical accident analysis: understanding the interactions between the task, individual, team and organization. In Safety in medicine, 2000, 101-116, Pergamon
137. Taylor-Adams, S, Vincent, C, Stanhope, N, Applying human factors methods to the investigation and analysis of clinical adverse events, 1999, Safety science, 31:143-159
138. Theaker, C., Mannan, M., Ives, N., Soni, N., 2000, Risk Factors for Pressure Sores in the Critically Ill, Anaesthesia, Vol. 55, PP. 221-224
139. Trautner, B., Darouiche, R., 2004, catheter-Associate Infections, Archives of Internal Medicine, Vol. 164, PP. 842-850

140. Trucco, P., Cagno, E., Ruggeri, F., Grande, O., A Bayesian belief network modeling of organizational factors in risk analysis: a case study in maritime transportation, 2008, Reliability engineering and system safety, 93: 823-834
141. Van der Gaag, L., Renooij, S., Aleman, C., Taal, B., 2002, probabilities for a Probabilistic network: A Case-study in Oesophageal Carcinoma, Artificial Intelligence in Medicine, Vol. 25, No. 2, PP. 123-148
142. Van der Schaaf, TW, Medical applications of industrial safety science, Quality and safety in healthcare, 2002, 11:205-206
143. Vincent, C, Risk, safety and the dark side of quality, 1997, British medical journal, 314:1775-1776
144. Vincent, C., Tylor-Adams, S., Stanhope, N., Framework for analyzing risk and safety in clinical medicine, 1998, British medical journal, 316: 1154-1157
145. Wang, C., Hybrid Causal logic methodology for risk assessment, 2007, PhD thesis, University of Maryland, College Park
146. Wann-Hansson, C., Hagell, p., Willman, A., 2008, Risk Factors and Prevention Among Patients with Hospital-Acquired and Pre-existing Pressure Ulcers in an Acute Care Hospital, Journal of Clinical Nursing, Vol. 17, PP. 1718-1727
147. Waterlow, J.,1985, Pressure Sores: A Risk Assessment Card, Nurs Times Vol.81, No.48, PP.49-55
148. Wick, J., Zanni, G., 2008, Skin Tears: Prevention and Treatment, The Consultant Pharmacist, Vol. 23, no. 7, PP. 508- 518

149. Wood, L., J. Ford, 1993, Structuring Interviews with Experts During Knowledge Elicitation, *International Journal of Intelligent Systems*, Vol. 8, No. 1, PP. 71-90
150. Wreathall, J, Nemeth, C, Assessing risk: the role of probabilistic risk assessment in patient safety improvement, 2004, *Quality and safety in healthcare*, 13:206-212